सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

२५५८४

I.A.R I.6.
GIP NLK—H 3 I.A.R.I.—10-5 55—15,000

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 43: 1948
NUMBERS 241–244

24988
IARI

Published Quarterly by the
AMERICAN STATISTICAL ASSOCIATION

WASHINGTON, D. C.

1948

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 43 MARCH 1948 NUMBER 241

ARTICLES

Statistics and Foreign Policy . . . . WILLARD L. THORP 1

History of the Uses of Modern Sampling Procedures . . . . FREDERICK F. STEPHAN 12

The Use of Mark Sensing in a Large Scale Testing Program . . . . WALTER L. DEEMER, JR. 40

A Proposed Basic Course in Statistics . . . . GEORGE W. SNEDECOR 53

Actuarial Estimates for Public Sickness Insurance Plans . . . . ABRAHAM M. NIESSEN 61

Earnings of Nonfarm Employees in the U.S., 1890–1946 . . . . STANLEY LEBERGOTT 74

Direct Determination of Compass Settings for Proportionate Area Pie-Charts . . . . HERMAN LASKEN 94

Profile Graphs. . . . . JOHN V. SPIELMANS 96

A Method for Obtaining and Analyzing Sensitivity Data . . . . W. J. DIXON AND A. M. MOOD 109

A Short-Cut Method of Fitting a Logistic Curve. . . . . WILLIAM A. SPURR AND DAVID R. ARNOLD 127

Statistical Methodology Index, No. 11 174

BOOK REVIEWS

ABAY, GYULA, *Bevezetés a Statisztika Tudományába: Rész I* . EUGENE LUKACS 135

ALBERT, ADRIAN A., *College Algebra* . . . . PAUL S. DWYER 136

BURTON, CLEMENT, *A Concise Manual of Statistics: With Special Reference to the Requirements of Students for Municipal Examination, Second Edition* . . . . HARRY PELLE HARTKEMEIER 137

DAHLBERG, GUNNAR, *Mathematische Erblichkeitsanalyse von Populationen* . . . . OYSTEIN ORE 139

DEWEY, EDWARD R., AND DAKIN, EDWIN F., *Cycles: The Science of Prediction* . . . . MILTON FRIEDMAN 139
. . . . MAX SASULY 141

FINNEY, D. J., *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve* . . . . MARGARET MERRELL 145
. . . . JOSEPH BERKSON 148

Guerreiro, Amaro D., *Manual de Estatistica* . . . Hugo Muench 153

Hays, Samuel, *An Outline of Statistics, Third Edition* Harold Nisselson 155
. . . . Edgar Z. Palmer 156

Heideingsfield, Myron S., and Blankenship, Albert B., *Market and Marketing Analysis* . . . . Lester R. Frankel 157
. . . . Philip J. McCarthy 160

Mortara, Giorgio, *O Custo de Produção do Homem Adulto e Sua Variação em Relação à Mortalidade* . . . . T. N. E. Greville 161

Schweng, Loránd, *Statisztika Módszertani Alapvetés* . Eugene Lukacs 135

Thompson, G. Clark, *Forecasting Sales* . . . . Charles F. Roos 163
. . . . H. D. Wolfe 165

Wright, Wilson, *Forcasting for Profit: A Technique for Business Management* . . . . William A. Spurr 167

. . . . Victor von Szeliski 168

Letters About Books . . . . John W. Tukey 172
. . . . John W. Dudley, Jr. 172
. . . . C. West Churchman 173

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 241 MARCH 1948 *Volume 43*

STATISTICS AND FOREIGN POLICY*

WILLARD L. THORP
Assistant Secretary of State

TO THE CARTOONIST, those who operate in the field of foreign policy fall into one of two categories. First are the black-haired, gaunt and cunning conspirators whose complicated and devious scheming is for objectives intended never to be apparent by methods which are sinister and super-top-secret. The other category includes the tea-drinking, spat-wearing butterflies who are too vacuous to have an objective and whose method is limited to charm, with the "r" relatively silent. Unfortunately, neither of these pen-and-ink conceptions even hints at the backbone of present-day foreign policy making and implementation, namely, careful theoretical and factual analysis by skilled technicians.

An increasing number of major problems of foreign policy are of such a character that measurement and magnitude become elements of basic importance. It is worthy of note that recently, when a delegation from the United Kingdom arrived to discuss with us a revision of the Bizonal Fusion Agreement for Germany, included in its impedimenta was a calculating machine. The classic "S" commodity list, beginning "shoes and ships," has been out of date for some time due to the lamentable obsolescence of that colorful item known as sealing wax, but the new "F" list—food, fuel, fibres, and fertilizer—has top priority on the current foreign policy agenda of most governments and these items appear there almost entirely as problems of magnitude. Today calories get more attention than kings. The slide rule, the calculating machine, and the statistical yearbook have become necessary tools of diplomacy.

The statistical puzzles arising out of foreign policy problems naturally are as varied as the universe. They are not limited by subject-matter, area, or time. What should be the number of whales permitted under the International Whaling Convention to be caught during the next whaling season in order to maintain a stable whale population, keeping also in mind the world shortage of fats and oils—a neat problem in the vital statistics and caloric content of whales? What are the proper cost-of-living-adjustment allowances to permit U. S. Government representatives in various countries to maintain approximately equivalent living standards—an index number problem with peculiar difficulties not merely for price level and foreign exchange reasons, but because of differing national customs of hospitality and patterns of protocol? What was the probable amount of destruction to American property in Italy arising out the War, and for which Italy must make partial payment under the Peace Treaty—a problem of statistical inference from quite fortuitous and incomplete sample data? What is the fair proportion between the United States and India for the shipment of raw cotton to Japan under present controlled trade conditions—a problem in which the accepted guiding principle requires the finding of a "representative base period" out of a most abnormal series of years? What is the relationship between the availability of tobacco products in the Ruhr and the production of coal and steel—a problem of psychological measurement since the proposal as made by several Senators rests in the allegation that tobacco products are particularly effective as incentive goods? What amount of goods sent to Russia under the Lend-Lease program was presumably unused and undestroyed at the end of the War and thus subject to a negotiated settlement—a problem in war-time and peace-time property life-tables, and attrition and depreciation rates? This random list of a few problems may serve to establish the inference of the presence of statistics in the State Department, but the record will be clearer if we consider two illustrations in somewhat fuller detail.

---

* Presidential address delivered at the 107th Annual Meeting of the American Statistical Association on December 29, 1947.

In a world where there are desperate shortages of commodities, the problem of allocation has become a matter of prime importance. Countries have become competing purchasers, and even, in a few tragic cases, competitors for relief assistance. The shortages are widespread and severe, and for many commodities there are few countries with an exportable surplus. Foodstuffs are in this category and no government with any claim to responsibility can look away while its people are hungry. The State Department has probably received more aide-memoires, notes, memoranda, and formal and informal visitations from Prime Ministers, Ambassadors, foreign technicians, and even self-appointed representatives, concerning the subject of food allocations than any other single topic during the last two years. The White House too has had distinguished callers on the same subject.

The development and application of the concept of equitable food allocation, based on a careful examination of requirements and availabilities, was done first by a small international committee, then by the International Emergency Food Council and is now in process of being taken over by the Food and Agriculture Organization, one of the specialized agencies of the United Nations. This international body makes recommendations to the supplying countries as to the proper distribution of their surpluses and these recommendations are followed with little variation.

The problem is a most complex one. The basic unit for comparing food levels is the calorie, but unfortunately the simple definition in Webster that a calorie is the amount of heat required to raise the temperature of one kilogram of water one degree Centigrade has not been so exact and indisputable when applied in the field of nutrition. Two caloric tables for valuing foods are in general use now, one by the U. S. Army and one by the International Emergency Food Council. There are at least half a dozen other tables used in various parts of the world. The two principal ones vary in caloric content from ten to fifteen per cent. Similarly, the effort to measure various types of grain in "wheat equivalent" opens the door to a thorough state of confusion. It is self-evident that when statisticians of many countries meet together to discuss a given problem, a primary requirement is that there be some common measure for setting down the facts and this has been a major task in the food field. Although this does eliminate one standard area for professional controversy, there will always remain enough other factors of disagreement to permit full self-expression.

Obviously, the first information required for making international allocations is that pertaining to the requirement and the indigenous supply in each country. If these can be satisfactorily determined, the import requirement follows merely by subtraction. At once it is necessary to remark that in many countries where industrial production is lagging the production of statistics is also below pre-war both in quality and quantity. Unfortunately, statistical organizations have been disorganized at the same time that the items to be measured have been subjected to wide variation. Even as basic a datum as population must be approached through estimation. Substantial movements of peoples have taken place, in addition to the abnormal effects of war on both birth and death rates. And none of the devastated countries has had the time or energy to take a post-war census, necessary to obtain a relatively secure benchmark. Similarly, agricultural production in most cases is measured much less accurately than before the war by the surviving statistical agencies in the various governments.

And even less certain, both abroad and at home, are the important estimates of wheat and coarse grains consumed by man and beast on the farm and thus not moving into the available supply. However, on the basis of the pre-war picture for which more reliable data are available, and the records for the past two years, including ration levels, the amounts imported, and the apparent stocks on hand, estimated import requirements for the current year, quarterly and by months, are worked out regularly by the experts.

On the other side of the equation, the possible export surplus, three of the chief exporting countries have regularly indicated their availabilities as best they can. Argentina, which provides about twenty per cent of the world's exports of grain, has not participated in the effort to plan the most effective distribution of the available supply. This unfortunate situation has been met in part by the other exporting countries through adjustment of the allocation to offset supplies obtained from the Argentine. Russia's exports, which have been limited in amount and have gone to very few recipient countries, have also not been subject to international allocation.

But even the direct facts on immediate requirements and supplies are not enough to solve the problem. Wheat, of course, is only one of many foods. It happens to be the cheapest form in which calories can be purchased in substantial quantities. In allocating wheat, the international committee must consider what other foods are available in the country with a wheat deficit. Obviously, it is quite proper in the light of the world shortage to cut down on wheat shipments to countries which have a fair amount of other foods. Even then, however, one must have some regard for the necessities of balanced diets. Furthermore, food habits and food requirements must be considered. Even in times of great need, a people does not change its way of eating over night. New foods are not easily introduced, even to a starving people, and established prejudices are surprisingly tenacious. For example, corn is not regarded as a proper food for human beings in a number of European countries, nor are potatoes in the normal diet of Italy, while rice is consumed much more than wheat in Cuba. Differences are not merely the result of taste. Countries with long winters have different requirements from those in the tropics. And even in our own experience as an occupying power, we have recognized that a much higher caloric requirement is needed for Germans than for Japanese to maintain the same health level. So the statistician, in figuring the amount of wheat to allocate to any deficit nation, must have a clear picture of that country's historical eating habits and be familiar with its preferences and prejudices.

Finally, in considering the amount which is to be permitted to come from abroad, the allocations must not work in such a way as to punish the country which brings its maximum to the market-place, or all enterprise and initiative in the direction of improved collections from the farms will be destroyed. Conversely, there must be some penalty for failure to use the indigenous supplies most efficiently.

Since the beginning of the allocation procedure, it has always been true that the screened requirements for foodstuffs for all the deficit countries have totaled to substantially more than the availabilities, and here the really painful job begins—the effort to determine where the requirements can be cut with the minimum of hardship. The figures of each country are reviewed again and again, and there are many conferences to explore various aspects of the situation more thoroughly. Finally, the allocations are announced. At least, the process has made everyone aware of the limitations on supply and the urgency of the demands from other countries.

These random comments about the allocation machinery may make the task appear exceedingly complex. But the fact remains that the job must be done. These formidable calculations, aimed to take into consideration both the overall requirements and supply situation, and the peculiar circumstances in each case, are the only hope of providing some basis of fairness and equity in the distribution of scarce things to people who are in desperate need of them.

There is no question but that living for millions of individuals for the next few years at least, will have to continue under rationing and allocations of critical, scarce commodities. The people in these countries know that death from starvation is just as permanent as death from bombing. They know that allocations and rationing are protections to their lives—that the rationing of milk, for instance, may cut down the number of fancy dishes served in fine hotels, but it does get the needed food to more mothers and infants for whom it is an essential. The international allocations in the same way are an effort more nearly to equalize the burden of the shortage on the people of the various countries, not leaving the distribution solely to ability to offer the highest bid or to the appeal of political sympathy, obligation or reward.

The decision having been made to disregard economic bargaining or political discrimination and to place allocations on an objective basis, the key to this whole process becomes the little-heralded statisticians—both those who must present the case for their countries, and those who must screen the competing claims and bring them into a reasonable relationship with each other. The day and night work and worry is theirs, but they have built in large part upon the work of other statisticians whose work had gone before. I hesitate to think how impossible it would be to handle this problem had it not been for the continued collection and analysis of agricultural statistics and nutritional data in many countries for many years.

As a second illustration may I speak briefly about the European Recovery Plan? The last six months have seen a most difficult and complex statistical undertaking in Washington—the examination of the requirements for European recovery and the study of the capacity of the American economy and other economies to carry the European deficit in the meantime. This task has absorbed the full energies and capacities of many experts in many government agencies. The only relief for the central group directing the project was temporary when in a lighter moment they decided to call themselves the Technical Wizards on the European Recovery Program, or the TWERPS, for short.

Anyone familiar with Washington during wartime can easily visualize the time and energy required to develop the details of a plan involving sixteen countries and a four-and-a-quarter year period. I remember in the late twenties being told by a Russian economist about the tremendous efforts required and the manpower devoted to drawing up the five-year plan. Last summer, I saw French economists and statisticians in a state of near exhaustion from working on the so-called Monnet Plan. No one should regard an undertaking of this kind lightly. There have been no days, and even at times no nights, of rest.

This project stems back, of course, to the suggestion by Secretary Marshall that the countries of Europe get together, examine what their requirements will be over a period of time sufficient to permit them to put their economies on a self-supporting basis, determine how much of these requirements they can meet separately and collectively, and thus indicate what additional help is needed to accomplish the program.

Sixteen of the Western European nations met in Paris last summer and, in an incredibly short period of weeks, drew up a program on which they could all agree. Undoubtedly, this agreement was possible because the essential elements of European recovery are beyond dispute—that production must be substantially increased, sound currencies must be established, and the restrictions on trade must be reduced. Their national requirements were presented and assembled. After some slight screening and reduction where the requirements were clearly beyond the possibility of supply, the so-called deficit was calculated. This was all incorporated in the Report of the Committee of European Economic Cooperation which, together with a number of technical annexes, was sent to Washington in September.

Work was already well under way in Washington by that time, particularly with reference to the capacity of our own economy to meet such foreign demands and the effect of such an operation upon our own economic operations. But the review of the European plan has proved to be a most complicated undertaking. Covering a period of four-and-one-quarter years, the program for each of sixteen countries and Western Germany had to be consistent as between its internal program and its export and import programs. For the total of all countries, the requirements from abroad and the availability of supply had to balance. Similarly, for the various individual commodities, demand and supply had to balance. For each country, its balance of payments had to be in equilibrium. And when supplies could not be obtained in the United States but could be found in other supplying countries, these prospective sources had to be determined and incorporated into the pattern. In other words, the total pattern had to balance not merely as to an overall figure, but by commodities, by country physical requirements, and by country balances of payments. Similar patterns had to be prepared for each year within the period. And finally, the attempt to achieve both commodity and the balance-of-payments estimates in turn had to be broken down by currency areas in order to indicate the nature of the deficit with the dollar area.

The number of arithmetical calculations which have gone into these estimates probably total more than a million. Five time periods are covered—the last quarter of the present fiscal year, April 1 to June 30, 1948, and the four successive years until June 30, 1952. Twenty-three areas were involved beginning with the sixteen countries which participated at Paris, the dependent areas of the United Kingdom, Belgium, France, Netherlands, Portugal, and Western Germany in three parts—the bizonal area, the French Zone excluding the Saar, and the Saar itself (since the Saar territory may shortly be incorporated economically into France). Twenty-six commodity groups were selected for particularly intensive treatment, of which a number such as iron and steel have to be built up from a series of separate categories—pig iron, scrap steel, iron ore, crude and semi-finished steel, tin plate, steel sheet and other finished steel—and for projections of the volume and value of trade covering movements of commodities among the participating countries, between the participating countries and the United States and other Western Hemisphere and other non-participating countries in turn. After the volume of this trade was derived, it had to be multiplied by prices to obtain values. While July 1, 1947 prices were used, it was necessary to ascertain prevailing prices in different areas of the world on that date since the prewar assumption that prices for internationally traded commodities tend to be equal the world over has lost its validity under the conditions which prevail today.

The task of combining the figures provided by the commodity committees into a coherent system from which balance-of-payments estimates could be derived fell, as it happened, to the Department of State. An early courageous attempt was made to grapple with it by assembling the adding and calculation machines of which the Department can boast only a sparse and scattered population, and by amassing at the same time the clerical assistance necessary to man or "woman" the machines. The attempt was futile. The traditions of the Department of State and its personnel training are oriented more toward the accurate and careful phrasing of a memorandum than the well-multiplied, checked and proven statistical table. Resort was necessarily had to punch cards, and automatic sorting and addition. It may be that the Foreign Office of 10 years hence will boast a full line of international business machines with operators in 24-hour attendance. As of today, it was necessary to work the calculations in on the graveyard shift at Census, Bureau of Labor Statistics, and finally in the Department of National Defense.

One further difficulty was to establish the price assumptions to be used for the future period. Here the crystal ball was particularly cloudy. The Paris Conference had used the prices then current, July 1, 1947, as the basis for both exports and imports for the first year (1948) and had then assumed that European export prices would remain firm, while import prices would decline by 7½ per cent in 1949, 10 per cent in 1950, and 12½ per cent in 1951. The American reviewers have felt that the only way out of this dilemma was to present a range. Actually, the basic calculations were made in July 1, 1947 prices, but the totals have been adjusted globally to meet different sets of price assumptions.

For the first fifteen months, all exports from Europe and imports to Europe from the Western Hemisphere except U. S., are calculated at 5 per cent above July 1, 1947 prices, while U. S. and non-Western Hemisphere shipments to Europe are 7.5 per cent above July 1, 1947. In the later years, the calculations on the high price assumption hold the price level constant with the level for the initial period, while the calculations for the low price assumption are based on a marked decline, particularly for the items imported to Europe. These different assumptions explain the presence of a range in the total requirement for the program of $15.1–$17.8 billions.

The projections are, of course, not blueprints which can be followed during operation. Actually, this is a sketch rather than a blueprint. Its purpose is to provide Congress and the public with as accurate an estimate as possible of what the program may in fact turn out to be, and the general magnitude of the requirements from abroad, if the European recovery program is to be accomplished. If some of the items are not available in the quantities indicated, there may be substitution. If availabilities increase, prices will fall, or if they decrease, prices will rise. In either event, the dollars involved will tend to be more nearly constant than the constituent elements. And with so many commodities and countries, we can fall back on the protection of all statisticians, the hope that the deviations will be somewhat compensatory.

In the original undertaking, five sets of questionnaires were drawn up by the European group to obtain information on food, fuels, machinery, iron and steel, transport, and balance of payments. This information has been available to us in Washington, and to it has been added further information which we requested, plus the vast reservoir of knowledge accumulated in our government. However, it is unfortunately clear that there are some serious gaps in the basic information required.

Even with the most complete information possible, there could be no assured results. The most that we can do is to achieve consistent and logical results from as reasonable assumptions as can be made. As I have already pointed out, assumptions had to be made as to price levels. Another uncertainty is created by the necessity to estimate crops. Should we assume that the weather will continue to be as uncooperative with the farmer in Europe as it has been since the end of the War? And there are many other unknowns, such as at what point the processes of commodity hoarding will cease and money will be used again as a store of value.

It is clear that the Recovery Program will have to be a dynamic and flexible operation. As was true during the War, programs will have to be changed from time to time as conditions change, both as to countries and as to commodities. To achieve the most effective use of the available resources, there will need to be continued and detailed statistical recording of progress made, and forecasts of the short and long-run prospects. The injection of statistical methods into foreign policy is therefore no temporary expedient—but promises to be a continuing necessity.

Certain conclusions are now apparent concerning our capacity to reduce the problems which I have been discussing to exact measurement. The first is the common complaint of the statistician—that we do not have adequate data. This is, of course particularly true of the countries where the effects of the War are still felt so severely. On the American side, we know much too little about the statistical quality and relevance of much of the foreign data. It is clear that these international projects require cooperation and understanding between the statisticians of all participating nations. And time can be used up most rapidly if one is skipping about among long tons, short tons, and metric tons, not to mention bushels, quintals, hundredweights, barrels and Imperial gallons.

On this point, that of the development of statistical data and the effort to achieve greater uniformity, there is much that the United Nations can do, supplemented by the private international statistical organizations. The establishment of a Statistical Commission by the United Nations and the international statistical meetings held in Washington last September may give us some encouragement. This is a long-time job—it calls for continuous support and stimulation. If much is to be accomplished, the statisticians in the United States must take the lead. We must continually be prepared to demonstrate that, in this modern world, many problems can be faced properly and solved economically only when measurement is respected as a fundamental characteristic of the analysis.

But beyond these points, there is a continuing frustration because too few relationships have been reduced to calculable form. It is obvious that planning really requires both cost and market data, that the requirements in the form of materials must be readily related to capacity, that labor supply and working capital requirements must all be part of such consideration. Here the statistician can make endless contributions, substituting detailed analysis for the rule of thumb or the experimental approach. In these matters, too much is at stake to be careless or casual. Nevertheless, our answers often contain far too much of the "rough estimate" and too little of the careful calculation.

The same problems arise in connection with the use of statistics in the foreign policy field as in any other—the struggle to do an honest and objective job, and the difficulties of convincing others that such is the result. There are always those who would prefer to be guided by emotion, prejudice, and preference. It is not at all surprising that suspicions which continually cross national boundaries are not stopped by statistics.

Nevertheless, the fact remains that problems must be handled;

there must be answers; and statisticians, both here and abroad, can contribute much to finding the best answers. As a professional group, we have a clear-cut responsibility. First, we must continually strive to improve our performance through developing our own capacities and improving the raw material with which we work. But beyond that, we must persistently try to persuade others to take an objective view of the facts, to use to the full the techniques and capacities which we have to offer. We cannot promise to solve all the world's problems. Many of them are not problems of measurement at all. But where possible the utilization of statistics, the reference to objective criteria, and the effort to measure before committing oneself to a line of action—these are all ways in which rational men approach problems. And foreign policy should be no exception to the rule.

Unfortunately, foreign policy is an area where it is all too easy for emotions and prejudices to be aroused, where problems all too often get subjective treatment. One approach towards international understanding is to expose problems to the facts whenever possible—and they must be accurate and dependable facts. The battle between prejudice and analysis is our battle, in which we must supply much of the ammunition. Our responsibilities and opportunities have greatly increased in the international field. We not only can help in the development of knowledge about the world we live in, but we can actually contribute substantially to that international understanding which is so greatly needed in the world today.

HISTORY OF THE USES OF MODERN SAMPLING PROCEDURES*

FREDERICK F. STEPHAN†
Princeton University

I. INTRODUCTION

MODERN sampling practice is the result of combining statistical theory with scientific knowledge about the material that is being sampled and with experience that has been gained through the use of various sampling techniques in surveys and experimental work. Its history stems from many roots and its applications branch out into many fields of science, commerce, manufacturing, agriculture, education, and government administration. In spite of its wide range of usefulness, sampling practice has been neglected in the training of statisticians, in the textbooks and treatises, and in the planning and analysis of most experiments and studies [1]. However, like Cinderella, it has risen from neglect to a position of well-deserved importance.

The history of sampling practice provides a useful background for the discussion of present theory and applications. Difficulties beset anyone who attempts to trace the developing uses of sampling because they are scattered throughout many branches of science and technology and are described, if at all, in subordinate portions of reports and articles whose titles provide no hint of what they may contain on the subject of sampling. Hence, this paper not only will touch each development very briefly and inadequately but no doubt it will miss entirely a number of highly important facts. (The author will appreciate suggestions for correction and addition.) The principal emphasis will be on the steps by which the present stage of progress in applied sampling was reached and no attempt will be made to present a comprehensive picture of present-day uses or to appraise the relative value of each contribution to the accumulation of technical knowledge.

There are two dangers that both author and reader must avoid in an historical review: (a) they may read into the records more than was actually there and (b) they may assume that nothing existed that is not given in the records. I hold to the view that the principles and technical practices of sampling emerged rather gradually from notions that were simpler, sometimes inaccurate and confused, and often not recognized by those who held them as being of any special importance. Hence the records are fragmentary and there were doubtless many instances in which principles were developed by two or more workers in ignorance of the other's work. I would not say that the same discovery was made "independently" for there is always a flow of common ideas, sometimes rapid and sometimes slow, in which all share. Hence it should be understood that the particular examples that will be mentioned are part of a larger development of practice and theory which connected them in various known and unknown ways and often influenced profoundly their characteristics and effectiveness.

---

* Presented at the 25th Session of the International Statistical Institute, a constituent part of the International Statistical Conferences, Washington, D. C., September 6-18, 1947. This paper will also appear in the Proceedings of the Conferences, which will be published in the near future.

† This paper was prepared in connection with studies of sampling under the Committee on the Measurement of Opinion, Attitudes and Consumer Wants of the National Research Council and Social Science Research Council.

Finally, greater attention will be given to the use of sampling in large-scale surveys rather than to its use in the laboratory or in small-scale (local) experiments.

II. EARLY EXAMPLES OF THE USE OF SAMPLES

The earliest examples of sampling procedures are to be found in certain very ordinary human activities. The common practice of taking a small part or portion for tasting or testing to determine the characteristics of the whole precedes recorded history and is one of the roots from which sampling methodology stems. Stirring and mixing before taking the sample is a prototype of randomization. The efforts of scientists and their predecessors to draw conclusions about the laws of nature from what they could observe in their immediate environment was a sampling process. Astronomy, perhaps the oldest of the sciences, started with the moon, the larger planets, and the stars that one could see from positions near the equator or north of it. Even today it is limited to those heavenly bodies that can be photographed with the best available instruments.

All empirical knowledge is, in a fundamental sense, derived from incomplete or imperfect observation and is, therefore, a sampling of experience. An unusually interesting example, not unrelated to modern population research, can be found in Halley's selection of mortality statistics in Breslau to form the basis of his life table. From this sample of one city he drew general conclusions pertaining to "the mortality of mankind" [2]. This was in 1693. Two centuries later, Sir John Lawes likewise used the annual record of wheat yields at Rothamsted on 5 plots totalling 33 acres to estimate the change in yield per acre from 1852 to 1879 for all of England and Wales [3]. While these examples were both statistical, all scientific observation, whether statistical or not, is based on sampling.

In spite of the widespread practice of sampling in commerce and industry and its utilization in scientific research, official statisticians have usually set as their goal the complete counting of a population and have sought to avoid any use of sampling or any inference beyond the bare description of that population at the time it was counted. Still they have been compelled to compromise, in the face of practical difficulties, on something less than complete coverage and often also on something less than adequate accuracy and quality in the information recorded. Such compromises have been common in the collection of many kinds of statistics, particularly those usually collected by voluntary reporting such as employment, wage rates, prices, and crop yields. It is also found in the collection of data that ought to be complete in order to satisfy certain non-statistical purposes, as for example in the registration of births. Often the official statistician, faced with great difficulties, opposition, and indifference on the part of the public and with only weak powers to support his efforts, could only compromise or resign his post. The result of the compromise was often a very crude and unsatisfactory form of sampling which strove to get as much information "as possible" and accepted what it got, instead of attempting to control the selection by some of the means that were available.

Interestingly enough, sampling has generally preceded the establishment of regular censuses. Prior to the first British Census in 1801 the size of the population of England was not known very accurately and various speculations were made and estimates prepared and disputed. Malthus projected his famous theory of population, and Adam Smith his *Wealth of Nations*, before the establishment of periodic national censuses. At least as early as 1754, estimates of population of England were made from the number of houses on the tax list plus a rude estimate of cottages not taxed, the total of dwellings being multiplied by a somewhat arbitrary factor of 6 persons per dwelling [4]. Other estimates were based on the reported number of baptisms, marriages and burials.

In 1800 Sir Frederick Morton Eden estimated the population of Great Britain at 9,000,000 using sample data on the average number of inhabitants per house as well as the number of births. The first census of Great Britain in 1801 confirmed his estimate. Somewhat similar methods of estimation have been used between censuses in the U.S.A. as recently as the last decade [5].

About 1765 Messance, and 1778 Moheau, published very carefully prepared estimates for France based on enumeration of population in certain districts and on the count of births, deaths and marriages as reported for the whole country. The districts from which the ratio of inhabitants to births was determined truly constituted a sample. Laplace prepared similar estimates in 1802 from enumeration of 30 departments of France and the reports of births, deaths, and marriages from 1799 to 1802. This followed a plan he published in 1786. He made a remarkable step forward in attempting to measure the precision of his estimate and announced that the odds were 1161 to 1 that it was not in error by more than 500,000 inhabitants or 12 per cent of the total population. Even though the method of estimate was crude and the measure of precision not wholly valid, Laplace's effort was much more successful than the complete census of France that was attempted at the same time. A similar procedure had been employed in 1784–91 by the famous chemist Lavoisier to estimate the number of horses, cattle, sheep, and pigs, as well as the area under cultivation [6].

The foregoing examples suggest that modern sampling procedure might have developed at least a century sooner than it did if it had received more attention from the scientists of the day. Only the officials of statistical bureaus and other posts continued to concern themselves with the problems of measuring population, and they were preoccupied with difficulties in the classification and interpretation of data, in the administration of statistical bureaus, and in the important problems of trade, finance, industry, agriculture, public health, etc. for which statistical data were needed. Hence they favored complete censuses or the closest approach to them that was feasible.

III. EXAMPLES OF DEVELOPMENT OF NEED FOR EFFICIENT SAMPLING METHODS

It would be interesting to explore the development of statistical work in all its many ramifications and inquire about the places in which modern sampling methods could have been used to good advantage. That is beyond the possibilities of a paper such as this. Instead examples will be given from the beginnings of four general lines of statistical work in which efficient sampling methods were needed: (a) agriculture crop estimates, (b) economic statistics of prices, wages, employment, etc. (c) statistical phases of social surveys and health studies, and (d) public opinion polling. The more recent developments of the use of sampling in these and other fields will be traced in succeeding sections.

(a) *Agricultural crop and livestock estimates:* At the present time a tremendous variety of statistics are collected in the United States on acreage planted in each principal or special crop, the estimated yield as judged at successive dates during the growing season, actual yields, numbers of livestock, equipment, farm labor, marketing, and other aspects of agricultural production. Other countries collect similar statistics with variations in scope, detail and frequency. In the U. S. as in many countries, the methods used for this purpose include periodic censuses, special surveys, voluntary reporting by selected respondents, and records produced in connection with taxation, marketing and foreign trade.

This vast system had its beginnings about a century ago. The first U. S. Census of Agriculture was taken in 1840. At the same time collection of agricultural statistics was begun by the Patent Office. Monthly reports were collected and published by the editor of the *American Agriculturist* in 1862 and in 1866 regular reports on acreage, condition of crops, yield, and livestock were begun in the newly established Department of Agriculture. Annual reports of prices were added in 1867. Thereafter the system developed into an extensive organization of agents and voluntary reporters that increased from 4000 in the late 70's to 300,000 in 1931 and a somewhat larger number now [7]. (The present system is described in a paper presented by Dr. C. F. Sarle at the International Statistical Conferences.) The methods of sampling were quite crude, many of the reports were simply based on the respondent's judgment about conditions in his locality and the list of respondents was built up of farmers, agricultural agents, members of the staff of agricultural colleges, etc. who were willing to serve without pay and were located at places widely scattered over the country. The system was a compromise between the excessive cost and slow returns of complete enumeration on the one hand and the wholly unsatisfactory alternative of relying on reports issued by private speculators or going without current information, on the other. Various attempts were made to improve the reliability of the statistics that resulted from these reports but they continued to be biased and of limited accuracy. Nevertheless they have been of great value to farmers and to other users of the statistics and they are the principal source of current data on agriculture. Here as elsewhere the introduction of modern sampling methods came late but is now making great contributions to the improvement of the statistical information in this field.

(b) *Economic statistics:* The history of economic statistics is a long and detailed story which can only be sketched here in inadequate outline [8]. Again, as for population and agriculture, there were early attempts to make the best of scanty and fragmentary material. Censuses and special investigations have provided increasingly valuable data but at long intervals and in inadequate detail.

In many countries, economic statistics derived primarily from the records produced by taxation and customs duties. This was true in the United States. Statistics about subjects that were not provided by the system of taxation had to be obtained by other means. More than a century ago the collection of reports of prices, wage rates, hours of work, employment, and production was begun with voluntary reporting by employers and occasional special surveys as the principal methods. Massachusetts established a statistical bureau in 1869, Pennsylvania in 1870, and other states at later dates; the Federal Government set up a Labor Bureau in 1884. These bureaus developed staffs of field investigators to collect data at monthly or other periods but relied heavily on mail questionnaires for many of their principal series. This continues to be the prevalent procedure although the establishment of social insurance systems and employment exchanges has provided other sources of statistics on earnings and employment.

Similar systems of collecting reports are operated by other agencies concerned with production, trade, and finance. These inquiries have been based on simple sampling procedures such as the use of a scattered group of voluntary reporters, except for certain relatively recent surveys. A serious effort was made, for example, to get wholesale prices "in representative markets" and to get labor data by sending agents into various districts with a list of employers from which they could choose what they believed to be a representative group, but the techniques of selecting a representative sample had not been developed as formal procedures. Here again modern sampling procedures have been introduced late but are making noteworthy contributions, in conjunction with other improvements, to the accuracy and value of these statistics.

(c) *Social surveys and health surveys:* The distinction between social surveys and certain related kinds of inquiries is difficult to make. Some, like LePlay's nomographic studies, are very intensive observations of a few individuals or families; others include thousands of cases but less intensively and with more attention to statistical analysis. Williams and Zimmerman listed 1500 studies of family living in the U. S. and elsewhere, made prior to 1935 [9]. Booth's Survey of London Life and Labour, Rowntree's survey of poverty in York, the reports of the Immigration Commission in 1907, and the Pittsburgh Survey of the Russell Sage Foundation are examples. While complete and thorough investigation of the problems was the ideal to which such surveys aspired, they were necessarily limited to a selected city, and to somewhat fragmentary data on pertinent questions by the sheer impossibility of a comprehensive canvass. Other surveys were made under the inspiration of their example but progress in methodology came slowly and as in many other investigations the accuracy of the results was seldom tested.

Sampling has been used extensively in studies of poverty and unemployment. During the depression of 1873–79 Carroll D. Wright used police in 19 cities and wrote to assessors in 375 towns throughout Massachusetts inquiring about the number of unemployed. His estimate of 28,508 corrected the current assertions that 200,000 or 300,000 workers were out of work [10]. Similarly in 1873 the number of unemployed in New York City was estimated on the basis of reports from 400 volunteer visitors of the Association for Improving the Condition of the Poor. These procedures were, of course, crude and inaccurate but they reflected a praiseworthy desire to measure what was being discussed in still less accurate terms by important public leaders. In 1893, unemployment was estimated indirectly by the decline in factory employment from its highest figure in a series derived from a sample of reporting employers constituting 70 per cent of the production in the state [11]. In 1893, estimates were collected by mail from respondents in 119 cities by Bradstreet. In New York City, Chicago and other cities, police and health inspectors surveyed samples of homes or factories [12]. In the depression of 1914–15, many surveys of unemployment were made, notably by canvass of the workers insured by the Metropolitan Life Insurance Co. through its agents in 29 cities [13] and by canvass of 104 city blocks and 3703 additional tenement houses by the Bureau of Labor Statistics using about 100 tenement house inspectors. Both agencies repeated their surveys in New York City eight months later [14].

In the depression of 1921–22, somewhat similar methods were used to estimate unemployment and, in general, while the results were far from satisfactory, there was improvement in the identification of the unemployed as distinct from other needy groups and increased detail in the information collected. Especially noteworthy as an advance beyond previous practice was the survey conducted in Columbus, Ohio, by Frederick E. Croxton and Mary Louise Mark using 100 university students to canvass three districts within the city selected by the presidents of the Federation of Labor and the Chamber of Commerce as a "fair sample of the wage-earning population of the city" and including approximately 10 per cent of the wage-earning population. The survey was repeated annually from 1922 to 1925 [15]. A similar procedure was used in Philadelphia by J. Frederic Dewhurst and Ernest A. Tupper who had school attendance officers canvass 166 small districts selected to provide an accurate cross-section of unemployment [16]. This survey was repeated annually during the 30's and provided several kinds of data not previously available. The Columbus procedure was used in 1929, 1930, and 1931 in Buffalo, N. Y. by Croxton [17] and in 1931 in Syracuse, N. Y. by John N. Webb [18].

Among numerous health surveys made in the U. S. from time to time, several examples may be mentioned that resemble the preceding surveys in their methods of sampling. In 1921, the U. S. Public Health Survey selected Hagerstown, Md., as a conveniently located and "fairly typical small city in the East not greatly influenced by immigration" in which to study sickness and fertility by periodic visits to a sample of households [19]. In 1929–30 the Committee on Costs of Medical Care made many sample studies, including one on the incidence of illness among 9000 families [20]. The U. S. Public Health Service and Milbank Memorial Fund studied sickness among those elements of the population that had borne the brunt of the depression by selecting severely effected districts in 10 localities, omitting the slums, Negro, and well-to-do sections [21].

(d) *Public opinion polls:* The practice of surveying public opinion emerged from simple but obscure beginnings in the "straw vote" conducted by newspapers to sense public reactions to candidates and obtain human interest stories by interviewing the "man in the street." Even before 1900 the *New York Herald* collected pre-election reports and estimates from all over the United States and attempted to forecast the outcome of elections. In 1904 it polled 30,000 registered voters in New York City. In 1905 the *Chicago American* and *Chicago Journal* took a poll during the mayoralty campaign and the *Columbus Dispatch* started a long series of polls by taking one in the state election of 1906. In 1912 papers in Boston, Chicago, Cincinnati, Denver, Los Angeles and St. Louis conducted a nation wide poll in 37 states. Four years later, this group of papers repeated its presidential election poll and another poll was taken by the Hearst newspapers. The Rexall drug store chain took a nation wide poll in 1920 using its stores as centers to collect the ballots. These and other efforts were outdone in size by *The Literary Digest* which started polling in 1916 and in 1920 mailed out 11,000,000 ballots to persons whose names were taken from telephone directories and other lists. Two years later it conducted a poll on prohibition. Other polls were taken at the rate of about one every year and a half. Robinson found that more than 86 straw polls were taken in the presidential campaign of 1928. Although these polls employed quite crude methods, they were successful in many instances because a moderate error only affected the prediction of the winning candidate when the election was close and because among many attempts some might be expected to turn out well. The sources of error in straw polls were analyzed in detail by Robinson in 1932 [22].

Closely related to the polls were market research studies and consumer surveys conducted by business concerns, publications, and advertising agencies. The results and methods of these studies received less publicity than the polls but they offered even greater opportunities for sampling than the straw votes. There was also a growing interest in public opinion research among political scientists, sociologists, and others which led to scientific interest in the improvement of the technique of opinion polling [23].

The foregoing discussion was limited to uses in which the sampling methods were quite simple, indeed often crude and inefficient, and not based on explicit considerations of probability. The dividing line is not sharp between them and some of the other surveys that will be included in later sections but they represent fairly well the pioneer stage in the emergence of modern sampling practice. This stage might also be illustrated by examples of experiments, surveys, and studies in other fields in which sampling in some form was used.

While, in these early instances, the sampling procedures were simple and usually employed uncritically with no great attention to accuracy and representativeness, it should be noted that the problems of observing and recording data were almost always far more serious than the problems of sampling. Modern sampling procedures, had they been available and had they been applied effectively, would have permitted the use of smaller samples and made possible in most cases more careful fieldwork. Through these developments and those in other fields of statistical study, there arose a great opportunity for the use of better sampling procedures that has been exploited only in part up to the present time.

IV. DEVELOPMENT OF THE USE OF THE SIMPLER RANDOM AND SYSTEMATIC SAMPLING TECHNIQUES

Much of modern sampling practice rests on processes of selecting individuals at random or according to certain systematic procedures. These processes may be employed as providing a simple method for selecting a sample or they may be used as parts of a more complex sampling scheme.

Random sampling utilizes the devices of games of chance or other methods that assure to each unit in the source from which the sample is drawn an equal probability before the draw that it will be included in the sample. It has the advantage of facilitating certain applications of the theory of probability to the analysis and use of the results.

Although the theory of probability was well established in the eighteenth century, its applications to the practical drawing of samples were delayed until the twentieth. Applications to games of chance and lotteries can be excluded from consideration since they were not concerned with estimating the characteristics of a population from a sample. Likewise, applications to the interpretation of data as having been produced by a random sampling process from a hypothetical universe can be excluded since they did not involve the actual practice of drawing samples. Also trials with dice or cards designed to "verify" known probability laws or merely for expository purposes are not examples of sampling practice. Doubtless there are many instances that have escaped the author's notice, but the earliest instances he has found are Bowley's sampling of a list of bonds and their interest rates in 1906 using the final digits in one of the tables in the *Nautical Almanac* to make the random selection and Student's testing of the t-distribution in 1907 by drawing cards from a large receptacle. Random sampling was facilitated by Tippett's tables of random numbers published in 1927 [24]. The practical importance of random selection slowly gained recognition, especially in connection with more complex procedures that were developed after 1920 and will be discussed in the next section.

Systematic sampling employs a simple rule of counting cases in some convenient order and selecting for the sample every nth case, or of using some similar pattern of selection such as taking material at measured intervals or drilling samples out of a large mass of metal by use of a template that locates the holes. More complex rules or patterns may be used and the starting point may be selected by a random choice or special rule. The theory of probability can also be applied to the results of systematic sampling under certain conditions. Bowley and others have interpreted it as a case of stratified random sampling and estimated sampling errors accordingly.

Systematic sampling has long been in use in taking samples of ore from the face of a vein in a mine or samples of metal from a pig or ingot. The practice of "cruising" a forest and sampling trees at uniform intervals has been well established in forestry for many years [25].

Systematic selection was utilized by A. N. Kiaer in his survey of Norwegian workers in 1895. As a means of facilitating special tabulations from census schedules, it was used in a study of family data in the 1900 Norwegian census, a study of marriage in the Danish census of 1901, and a housing study for Oslo in 1913–14 [26].

Kiaer's work apparently had little influence on surveys in the U. S. or in Europe. Apart from the discussions he initiated in the *Bulletin* of the International Statistical Institute from 1896 to 1903, it was mentioned only briefly by Edgeworth in his Presidential address before the Royal Statistical Society in 1912 and in the U. S. by W. B. Bailey and F. S. Chapin in 1906 and 1920 respectively [27]. It had a greater effect in certain continental surveys [28].

In 1909 Bowley included a brief chapter on Sampling in his Elementary Manual of Statistics making general references to sampling in commerce, mining, and industry, but not to specific sample surveys. However, in 1912 he took every twentieth working class household in Reading in a study of poverty and his associates did the same for three other cities in 1913 [29]. He computed probable errors of sampling and recognized other sources of error in the interviewing, definitions, and process of estimating. This study was followed by a systematic sampling of census schedules in 1915, and by similar surveys of Liverpool (1930) and Merseyside (1931) by Caradog Jones, Southhampton (1927f) by Ford, and London (1929) by the London School of Economics under Bowley's direction [30].

John Hilton made a series of studies of workers in the unemployment insurance system beginning in 1923 [31]. They were selected systematically from the files of the Labour Exchanges. In addition to the information in the records some data were obtained by interviewing workers who came to the exchange. Certain deviations from strictly systematic selection introduced biases which were reduced by subsequent improvements of the method. Hilton found a sample of only one per cent quite satisfactory to meet the practical administrative and policy-making purposes for which the studies were made. The reduction in expense that resulted from sampling such a small proportion of the records was indeed impressive. Oddly enough the method was not imitated to any great extent by other government bureaus.

Another example of the great usefulness of sampling to get quick inexpensive tabulations was furnished by the Japanese who took a sample of one in 1000 from their 1920 census to get data quickly after the earthquake of 1923 [32]. Subsequent checking when the usual tabulations were completed verified the accuracy of the sample results. Further examples of sampling census schedules for special tabulations are Day Monroe's study of Chicago families in the 1920 Census and R. F. George's study of British workers in the 1931 Census [33].

In spite of this successful experience, the use of random and systematic sampling procedures in statistical work made slow progress. At the Rome Session of the International Statistical Institute in 1925 a resolution was adopted recommending the use of sampling for statistical purposes with appropriate precautions as to its representativeness, mathematical statement of the precision, and full description of the methods employed. The reports that were submitted by the commission that drafted the resolution presented evidence of the usefulness of sound sampling procedures but both the reports and the resolution had less immediate effect on statistical practices than might have been expected. Eight years later, in 1934, Neyman revived the discussion of these reports in a notable paper before the Royal Statistical Society.

In America, serious attention was given to problems of sampling methodology by several committees of the Social Science Research Council. There was a strong current of interest among the sociologists including Chaddock, Ross, Ogburn, Stouffer, Lundberg, Dorothy Thomas, Stephan and others [34]. Margaret H. Hogg, who had worked under Bowley's direction on some of the British surveys, came to America to the staff of the Russell Sage Foundation and there made a critical study of employment and unemployment statistics. In an article in the *Journal of the American Statistical Association* she made a strong plea for rigorous methods of sampling and cast doubt on the value of surveys such as those that had been made in Philadelphia and Buffalo, in which the sample was selected by judgment rather than random procedures. She was equally concerned about the adequacy of the classification of workers into the various categories of employment and unemployment and the analysis of the data in terms of significant questions. In the spring of 1931 Miss Hogg made a survey of unemployment in New Haven, Connecticut, partly for the purpose of testing the practical difficulties of applying a random sampling method and also developing better schedules and statistical categories for unemployment surveys [35].

Two very important general developments affecting the use of sampling in the United States occurred in 1933. One was organization of large-scale work projects for the unemployed under the national programs of the Federal Emergency Relief Administration (1933–35), Civil Works Administration (1934–35), and Work Projects Administration (1935–40). The second was the enlistment of many leading statisticians from the universities and business in the reorganization of government statistical work and administration of emergency agencies [36]. The statistical needs of the government increased tremendously as it took active steps to meet the problems of the depression and undertook various New Deal programs. Since the American statistical system is decentralized, it became necessary to establish a Central Statistical Board to coordinate and regulate the statistical activities of many agencies engaged in collecting data for their own purposes or for general use. The Board was largely an outgrowth of the Committee on Government Statistics and Information Services, an advisory group organized by the Social Science Research Council and the American Statistical Association at the request of the Secretaries of Agriculture, Commerce, Interior and Labor to assist them in overhauling the statistical work of their Departments.

The great variety of statistical work done by the regular agencies, the National Recovery Administration, and other emergency agencies far exceeds the scope of this paper. Only a few large scale surveys can be described briefly as examples of those of their projects that were based on sampling; many smaller and more specialized studies were also undertaken. A number of difficulties prevented these projects from embodying fully the best sampling procedures then known and from exhibiting the range of uses to which sampling methods could be put. Most of the work was started with inadequate time for planning and preliminary trials. There were, curiously enough in a period of unemployment, severe shortages of persons technically competent to supervise the work. Moreover the field workers and office staffs were drawn largely from the unemployment relief lists and worked under restrictions as to hours and rates of pay. In spite of these difficulties the results contributed greatly to many very important programs and provided detailed statistical information where there had been little or none before.

(1) The Financial Survey of Urban Housing was taken in conjunction with the Real Property Inventory, a complete canvass of housing in May 1934 covering 60 cities. The Financial Survey was taken by interviewers who visited all the families in one block out of every ten in larger cities and one in seven in smaller cities. Additional schedules were obtained by mail from families in another four blocks out of ten. The survey resulted in a 1200 page report of detailed financial data and was unique as an early instance of the use of sampling procedure

for obtaining special information as part of a single complete canvass [37].

(2) A series of research studies of workers on relief and related subjects were undertaken by the FERA and WPA in selected cities and counties from 1933 to 1940 using a variety of methods and samples. The localities were selected in part by judgment of their representativeness of different types of situations that were to be compared, in part by administrative considerations of availability of supervisors, noninterference with other projects, etc. Systematic sampling was used extensively within the localities for many of these studies and the samples were appraised critically in the reports. One of the studies was an 11 per cent national sampling of workers on relief in March 1935, with proportions varying from state to state, providing very detailed tabulations by age, occupation, and education. The accuracy of the sample results was verified by comparison with subsequent complete tabulations and a table of sampling errors was given [38].

(3) The Study of Consumer Purchases was conducted in 1935–36 by the Bureau of Labor Statistics in 32 cities and the Bureau of Home Economics in 19 small cities and 206 villages or rural counties [39]. The study was a work relief project and other agencies participated in the planning and analysis. In selecting the localities, available data on their economic characteristics were used to achieve a high degree of representativeness; within each locality families were selected systematically from city directories, the schedules of the Real Property Inventory (New York City), the 1934 WPA Census of Chicago, and similar lists. The smaller areas were canvassed completely. From the first simple schedules obtained for 783,000 families a second sample was selected for interviewing about details of income and expenditures. This sample was designed to provide adequate numbers of families in certain relatively rare categories without taking more families in the more prevalent categories than were needed for comparisons between the various types and kinds. The sampling procedure was complicated by the necessity of interviewing in such a manner that if the work were terminated at any stage in the process families that had been visited would still be a representative sample. This project far exceeded any previous study of its kind in size and complexity. The procedures and problems have been described rather fully in the reports and elsewhere. They follow a basic plan that was prepared by the Social Science Research Council in 1929 [40].

(4) The National Health Survey was conducted by the U. S. Public Health Service in 1935–36. It covered 83 cities in 18 states with sampling

ratios varying from 1 in 38 for New York City to complete canvass of the smaller places. The sampling was done by dividing each city into small areas of less than 1000 persons each using the 1930 Census enumerators' maps as the basis, and then drawing a systematic sample of these small districts for complete canvass. The Survey supplied a great amount of data on illness and medical care. Its schedules were also tabulated in great detail by the Social Security Board to obtain cross classified information about family composition and characteristics for actuarial studies.

(5) The Michigan Census of Population and Unemployment was taken early in 1935 by the State Emergency Welfare Relief Commission as a WPA work relief project [41]. The sampling procedure involved complete enumeration of all cities between 3000 and 40,000 inhabitants and a 20 per cent random sample of smaller places and rural townships. Cities of more than 40,000 inhabitants were sampled by taking all dwellings with certain predesignated house numbers according to a scheme developed by S. A. Stouffer. The reports not only provided detailed data on unemployment in various parts of the state, but also unusual information on migration that was subsequently analyzed by John N. Webb and Albert Westefeld [42].

(6) The Minnesota Income Study was a work-relief project sponsored by the Minnesota Resources Commission and operated in 1938 and 1939 using a complete system of stratified sampling with a systematic selection of households in urban areas and a random selection of two-square-mile areas in rural districts [43].

(7) The New York City Youth Survey was made by the Welfare Council of New York City in 1935 taking every hundredth household from the lists of the 1934 Real Property Inventory supplemented by a similar sample of dwellings built after the Inventory [44]. It was also a WPA work relief project. All members of these households between 16 and 24 years of age were interviewed about their leisure activities, employment, schooling, and other personal data.

(8) The Continuous Work History Sample is a 4 per cent sample of workers covered by old age insurance maintained by the Social Security Board in connection with its records of earnings and contributions to the insurance scheme [45]. Beginning with 1937 all reported earnings and related information are compiled for this group of workers. They were selected by use of certain combinations of digits in their account numbers. The sample provides annually distributions by age, sex, race and taxable wages that are needed for actuarial purposes as well as useful for general economic studies.

(9) The 1937 Enumerative Check Census of Unemployment was conducted by the Director of the National Unemployment Census who had been ordered by an act of Congress to undertake a voluntary census of the unemployed by mail [46]. The Check Census was designed to determine what proportion of the unemployed had actually registered and thereby provide more accurate estimates of the volume of unemployment than had been available previously. The sample was a systematic selection of all households in 2 per cent of the postal carrier routes in the U.S.A. About 18 per cent of the population, not served by carrier delivery, was excluded but was taken into account in the estimates. The 510,000 households in the sample were visited by postal carriers to obtain data of employment and unemployment and a special count was made of the number of workers and unemployed persons for the routes that constituted the sample. From these data it was determined that 71 per cent of the unemployed had registered and estimates were prepared for separate classes of workers with corresponding calculations of the sampling errors of the estimates. The analysis of the estimates and their errors was especially noteworthy.

The foregoing examples exhibit progress in utilizing sampling technique for statistical surveys of exceedingly important national problems. They reflect the trend toward greater use of detailed statistical data and research as a basis for determining policies and establishing programs on a city and state wide basis as well as nationally. They are impressive for the size of the samples taken, but there were also many smaller samples of comparable importance. In two ways, they were a result of the depression: (a) in the urgent demand for statistical data generated by the emergency and by the New Deal programs, and (b) in the funds and relief personnel that were made available. However, these influences tended toward haste and inadequate preparation, toward the use of less efficient sampling schemes in order to meet problems of supervision and the restrictions on expenditures for travel by interviewers, and toward types of inquiries that were suitable for the kinds of interviewers that were available. Problems of cost and economical operation were considered but were affected by the primary purpose of providing useful work for unemployed white collar workers. This tended to discourage intensive analysis or research studies not of a relatively routine character since the amount of employment provided in relation to the kind and amount of supervision required was smaller for such projects. Finally sampling problems received only secondary emphasis because the other problems connected with the surveys were much more serious and time consuming. The results of these surveys were of very great value and to a large extent this was due to the use of sampling procedures.

Coupled with the direct effect of depression problems in stimulating the use of sampling was the trend of thinking among statisticians and social scientists who, during the late 20's and early 30's, became increasingly interested in applying error theory to the analysis of time series and other data [47]. This trend of interest was well represented among many of the men who participated in the work of the Committee on Government Statistics, and recommended in 1933–35 greater and better use of sampling procedures in government statistical work for such purposes as (1) checking the coverage, accuracy, and completeness of censuses, (2) testing the representativeness of current reporting systems (crops, cost of living, payrolls, employment, production, etc.), (3) supplementing the regular schedules in censuses to relieve them of overloading, (4) obtaining prompt national and regional figures on such subjects as unemployment, and (5) making very detailed intensive studies in which the emphasis is on the analysis of relationships [48]. The Committee also recommended that the staff of the Central Statistical Board and the Bureaus of Agricultural Economics, Census, and Labor Statistics include expert advisers on sampling procedure whose formal professional training in the technical phases of sampling was comparable to that of chemists, physicists or biologists in the government service. These agencies did move to accomplish what had been recommended. They experimented with trial surveys of unemployment, agriculture, construction, retail prices, and basic studies of the problems involved in developing practical sampling systems [49].

By this time sampling practice had outgrown the simpler methods of random and systematic sampling and was developing complex sampling systems specially designed to fit the nature of the population that was being sampled, the costs and administrative factors, and the principles of efficient design that had evolved in scientific experimentation and industrial production.

V. EXPERIMENTAL COMPARISON OF METHODS AND DEVELOPMENT OF COMPLEX SAMPLING PROCEDURES

Modern sampling practice utilizes a variety of devices and methods which have been tested experimentally and incorporated into sampling theory. The aim of modern practice is to select those methods which, when combined in the most appropriate manner, will constitute a sampling system that is economical, convenient, and accurate in the situation for which it was designed. Among the devices available for the construction of such a sampling system are: (a) the technique of selecting items at random, (b) the determination of the kind of sampling unit that will have the most desirable properties, (c) subdivision or stratification of the population, in an advantageous manner, (d) the use of variable sampling proportions, (e) subsampling and multistage sampling in which the sample is itself sampled in turn, (f) drawing two or more units from each ultimate subdivision to permit estimation of the sampling error from their differences, (g) use of information provided by variables that are correlated with the one that is being studied, and many other procedures.

It is difficult to find the original source of many of these techniques but it is clear that they were, to a large degree, crystallizations of notions suggested by common sense, customary practices, and practical problems encountered in sampling. Discovery of their importance and success in combining them into efficient sampling schemes has required a clear understanding of their properties and a considerable amount of ingenuity. Beyond that, careful experimentation has been necessary to test their relative merits under various conditions. Their development has occurred primarily in connection with agricultural experimentation, mass production industry, and large-scale surveys.

(a) *Agricultural experimentation:* One of the first steps in the experimental comparison of sampling schemes was the study of heterogeneity in the fertility of a field and the consequent correlation of adjacent portions. In 1910 and 1911, very important work was reported by "Student," Wood and Stratton, and Mercer and Hall [50], leading to the conclusions that in field trials there is always some experimental error, that it can be reduced by taking a large number of small plots and, that, in the comparison of two varieties or treatments, it can be reduced by taking these plots in pairs, adjacent to each other, and using one plot out of each pair for each variety or treatment. The latter device was adopted so as to take advantage of the correlation between adjacent areas. Plots of different sizes and shapes were compared using the results of a careful harvesting and measurement of mangolds and data on 400 reports of duplicate plots reported by other investigators. These results led to similar experiments in England, America, and elsewhere [51]. Various systematic sampling schemes were proposed and tried experimentally, both for locating plots and subsampling the plots in order to reduce the labor of harvesting them completely. In 1929, Clapham and Wishart published the results of experimental comparisons between certain random and systematic methods of sampling a field of potatoes or cereals and analyzed the factors that contributed to the error of the estimate of yield [52]. In the same year Smith and Prentice published results of a field study in which they took samples of soil and then subsampled for laboratory examination. In their analysis of the results they estimated the increase in error due to subsampling [53]. R. A. Fisher, whose advice and methods of statistical analysis had been followed in many of the experiments in the 20's and later, developed complex systems of experimentation of great efficiency and revolutionized experimentation [54]. Yates and Zacopanay published in 1935 a thoroughgoing analysis of the efficiency of sampling a field, including comparative labor costs [55]. A number of manuals were published expounding the principles and practice of field experimentation. This work of the agronomists and statisticians had direct applications to large-scale surveys.

(b) *Mass production industry:* Another source of complex sampling procedures was the development of inspection and quality control in mass production. In 1923, engineers at the Western Electric Company, manufacturers of telephone equipment, began to apply probability theory to the inspection of telephone exchange equipment. In the refinement of manufacturing processes and progressive adoption of more rigorous specifications, the cost of inspection had become very burdensome and sampling procedures offered possibilities of reducing the amount of inspection without serious loss of control over the quality of the product. Different sampling schemes were compared on the basis of their probability-of-acceptance curves and tolerances were established for the proportion of defective pieces in a lot. Thereafter sampling procedures were developed in more complex forms to achieve a reduction in the amount of inspection to the minimum necessary under the conditions of manufacture and the specifications to be met [56]. Technical problems more or less peculiar to the telephone industry made it the pioneer in the applications of mathematical statistics to industrial processes, including not only inspection but quality control and probability analysis of operating problems under Rorty (as early as 1903), Shewhart, Dodge, Molina, Fry, and others [57].

(c) *Communication of new techniques:* Those developments in agriculture and engineering had both direct and indirect effects on sampling survey practice. They provided principles of design and contributed to the growth of applied mathematical statistics. Still there were many practical problems and obstacles that delayed the immediate extension of the methods developed for field trials and manufacturing to large-scale surveys. One of the obstacles was the relative lack of communication between statisticians engaged in different types of work. An important step toward facilitating communication was taken by the Royal Statistical Society in 1933 when it formed the Industrial and Agricultural Research Section and published the proceedings of meetings of the Section as a *Supplement* to the Society's *Journal*. The Section brought together mathematical statisticians, engineers, agriculturists and others interested in practical applications of newly developing statistical theory and stimulated greatly interest in sampling. E. S. Pearson presented a notable paper on "Problems of Industrial Sampling" at a meeting of the Section early in 1934 [58]. A few months later, Jerzy Neyman read before the Society a comprehensive theoretical and critical paper on stratified sampling and purposive selection. He cited among several examples a survey of the working population in Poland for which he had designed a stratified random sample of districts [59]. A number of other papers by various authors, published from time to time in the *Supplement*, included discussions of sampling practice.

In this period interest in sampling practice was heightened in Great Britain by Shewhart's visit in 1932 (a factor in the formation of the Industrial and Agricultural Research Section) and in America by the visits of Fisher, Neyman, Yates, and Cochran (1936–38) as well as by American students who studied in England.

(d) *Large-scale surveys:* Another root of complex sampling practice is to be found in the progressive improvement of large-scale surveys of crop yields and acreages, the labor force, and other economic and social facts. A pioneering survey in India in 1923 by J. A. Hubback went almost unnoticed until Mahalanobis discovered and republished it in 1946 [60]. Hubback undertook to estimate the yield of rice by objective measurements in the fields instead of by the methods based on the judgment of farmers and brokers. He used as sampling units very small areas measured off by a wooden frame in every field his agents found being harvested as they walked a fixed route from each selected center. Hubback was acquainted with "Student's" and Bowley's sampling work. His report emphasized the advantage of random selection of many very small areas as economical and necessary to obtain a valid estimate of error.

In the period 1928–31, J. O. Irwin examined the British crop estimating system and made recommendations for improving the sampling methods [61]. Subsequent work was done by Cochran. In the United States, W. F. Callander, J. B. Shepard, C. F. Sarle and others began to plan in the late 20's for a sample of farms to be used for various studies of acreage, production, livestock, etc. Sarle made comprehensive appraisals of the methods used to obtain farm-price data in 1927 and crop-yield estimates in 1932 and proposed a program of studies aimed at testing and improving the various sampling operations then in use [62]. Several studies of this type were made by state agricultural statisticians.

In July, 1936, the Bureau of Agricultural Economics and Iowa State College held a conference on sampling attended by many of the state agricultural statisticians as well as staff from Washington and Ames [63]. R. A. Fisher who was lecturing at Ames participated in the discussions. The conference set forth a great variety of practical problems and outlined specific projects for research to help solve them. Some of these probjects were started at Ames in cooperation with the Bureau, some were undertaken in Washington. The program has been continued up to the present time with additions and modifications. It has resulted in a remarkable series of reports and developments of methodology [64] as well as the establishment of a sampling staff within the Iowa State College Statistical Laboratory at Ames under G. W. Snedecor, A. J. King, and R. J. Jessen.

During the same period, P. C. Mahalanobis and his associates at the Indian Statistical Institute in Calcutta developed large-scale sampling methods to a high degree of technical perfection and utilized them for surveys of jute acreage, agricultural finance, famine conditions, radio listening, and many other subjects. The theoretical system associated with these practical applications has been set forth fully by Mahalanobis in his publications and lectures [65]. Mahalanobis has made careful experimental measurements of the human errors in sampling surveys by use of "interpenetrating samples" and other means.

From 1936 on the Research Division of the U. S. Bureau of the Census under C. L. Dedrick made extensive studies of problems in population sampling, utilizing the 1930 Census returns. The results have not been published. In 1939, Stephan, Deming, and Hansen designed a sampling scheme that was used in the 1940 Population Census to draw a 5 per cent sample totalling 7 million people. This method was based on experimental tabulations of 1930 census data and on a variation of systematic sampling [66]. The sample was used to obtain preliminary estimates of regular census statistics as well as tabulations of the special questions that were asked only of the people in the sample. Subsequent complete tabulations verified the accuracy of the estimates.

During the same period experiments were conducted in a number of cities and towns by the Research Division of the Work Projects Administration, under H. B. Myers and J. N. Webb, to test out methods of conducting sample surveys of unemployment. On the basis of these trials a monthly nation-wide survey was started in December 1939 utilizing a subsampling procedure in approximately 60 counties. Its scope was enlarged during the war to include housing and labor questions and in 1942 it was transferred to the Census Bureau where it was developed further as the Monthly Survey of the Labor Force [67].

With the Monthly Survey and other sampling surveys required by the war agencies, the Census Bureau had a sizeable program of sampling work and developed, under Morris H. Hansen, a staff of sampling specialists who designed sample schemes, and studied their performance in practice. Data from the 1940 Census for very small areas were made available for use in sampling and Sanborn maps showing every street and building in a large part of the cities in the United States were obtained [68]. In addition, airplane maps were obtained for a large part of the United States and procedures were developed for employing them in sampling.

A further step was taken in the development of sampling facilities when the Bureau of Agricultural Economics established at Iowa State College in 1943 a project to design and prepare a Master Sample of farms for use in its surveys [69]. This Master Sample is a stratified selection of very small areas based on airplane photographs and highway maps. The Census Bureau joined in its development and had it extended to include other than farm areas.

These most recent developments in complex sampling procedures have been described very inadequately and many other recent developments have been omitted for want of space. Other speakers on the program of the International Statistical Conferences will present the current stage of theory and examples of present day practice.

VI. DEVELOPMENT OF USE OF JUDGMENT SELECTION AND CONTROLS IN SAMPLING

All attempts to obtain data by sampling involve some choices based on the purposes of the study, knowledge of the material, and experience with the methods of observing and recording similar data in previous studies. For example Bowley's choice of Reading for his survey in 1912 and Clapham's choice of certain fields at Rothamsted for his experiments on sampling design were not random selections. Yet the conclusions drawn from such samples as these are commonly generalized to apply to other periods of time and other places. Some investigators have depended entirely on their knowledge and judgment to make the final selection of the items that constitute the sample, others have insisted that judgment should only be used to subdivide the population into strata and units of sampling, the actual choice of the sample being made at random. These views have been presented in various forms and the differences have not yet been fully resolved. Research is in progress to determine some of the facts and develop the principles that are involved in appraising the performance of various sampling procedures based on these conflicting views.

Jensen, in his report in 1926, described a number of surveys made by "purposive" sampling in which districts were selected which, taken together, matched the population in certain characteristics known from a previous census. The theory of purposive selection was developed by Bowley in 1926, and discussed by Jensen in 1928 [70]. Neyman compared purposive and stratified random sampling in 1934 giving special attention to experience reported by Gini and Galvani. Mangus described a variety of purposive sampling of counties used in studies of the FERA in 1934 [71].

The development of public opinion polling and market research involved many interesting problems of social psychology and survey procedure that fall outside the scope of this paper, but it also led to the use of many different sampling schemes, prominent among which were those termed "quota control" sampling. The earliest work was done with samples chosen to meet the convenience and very limited resources of the research personnel. In this respect they were like most other surveys conducted at the time. Robinson has reviewed in detail the history of straw votes and opinion polls to 1932 [72]. He also analyzed *The Literary Digest* poll which, in spite of its great size was badly biased by its method of mail balloting. Several analysts were able to improve *The Literary Digest* results by relatively simple adjustments, [73] but, in field surveys in which interviewers are left free to use their judgment in selecting a representative group of respondents, a better procedure appeared to be to regulate the selection of the sample by assigning to each interviewer a quota of persons of various kinds so as to obtain the correct proportions of men and women, old and young, rich and poor, etc. in each city or district that was canvassed. This procedure was developed by Cherington, Roper, Gallup, Crossley, whose sampling surveys of opinion became widely known after the election of 1936, and by market research and opinion survey organizations, whose studies received less publicity [74]. The general principles of these and related methods have been discussed by Brown, Franzen and Robinson, by Cassady, and in two symposia [75]. Quota methods have been employed in the surveys of the National Opinion Research Center [76], the Office of War Information and the British Wartime Social Survey

[77]. Most of the opinion polls and market research studies around the world at present employ quota methods or related procedures of sampling. These methods have been adopted partly by imitation of the more widely known surveys and partly because they are attractive with respect to costs and practical convenience.

VII. GENERAL VIEW OF THE PRESENT PERIOD

Current uses of sampling will be described by other speakers at this session. For the most part these uses were connected with the war which proved to be an even greater accelerating stimulus than the depression of the 30's. They included surveys of employment, housing, prices, shortages of consumer goods, dealer inventories, public reaction to wartime measures, characteristics of selectees, the attitudes of soldiers, effects of bombing in England, Germany and Japan, public opinion in many countries, inspection of election registers in Greece, surveys of saving and spending, employee attitudes, reading and radio listening, social and psychological factors affecting fertility in the U. S. and in Britain, etc. In manufacturing and procurement, sampling methods were perfected for quality control and inspection of materials, parts, and finished products. Market research and opinion studies by private research agencies and business concerns developed rapidly, with considerable emphasis on measuring public relations and employee attitudes. The volume of work increased and many continuing or repetitive surveys were established, making it possible to spread the costs and hence devote greater resources to the design of each sample than had been possible for infrequent sampling surveys.

Other speakers will describe the more recent developments that grew out of these earlier accomplishments. The *Journal of the Royal Statistical Society* and its *Supplement*, the *Journal of the American Statistical Association*, the *Annals of Mathematical Statistics* and many of the scientific and professional publications in various fields have published the increasingly numerous articles that reflect the progress of applied sampling. The reader should review them, beginning in 1940 or earlier, and search articles that report particular studies as well as those that are principally devoted to discussions of method.

Great as these advances have been, they are only the forerunners of a broad expansion of sampling in the future. Further improvements in method are also to be expected. Much experimental work is being done within and outside government bureaus. The Committee on the Measurement of Opinion, Attitudes and Consumer Wants, appointed jointly by the National Research Council and the Social Science Research

Council, is making a comprehensive study of sampling methods as they are used in its field. The United Nations Statistical Commission has established a sub-commission on sampling which will contribute greatly to the development of sampling surveys in many countries. The International Statistical Institute can perform a great function in extending the use of sampling procedures in the many statistical activities throughout which it exerts its leadership and influence.

NOTES AND REFERENCES

The following abbreviations will be used:

J. Agr. Sci.	*Journal of Agricultural Science*
J.A.S.A.	*Journal of the American Statistical Association*
J.R.S.S.	*Journal of the Royal Statistical Society*
Phil. Trans.	*Philosophical Transactions of the Royal Society of London*

[1] Even in the comprehensive *Handbook of Mathematical Statistics* (Houghton Mifflin, 1924) although considerable space was given to time series none was given to the procedures of drawing samples.

As an example of the neglect of sampling in one among many fields of science, see: Quinn McNemar's critical review, "Sampling in Psychological Research," *Psychological Bulletin*, 37, 1940, 331–365.

[2] Edmund Halley, "An Estimate of the Degrees of the Mortality of Mankind Drawn from Curious Tables of the Births and Funerals at the City of Breslau," *Phil. Trans.*, 1693, 596–610.

[3] J. B. Lawes and J. H. Gilbert, "On the Home Produce, Imports, Consumption, and Price of Wheat," *J.R.S.S.*, 43, 1880, 313 f.

[4] E. C. K. Gonner, "The Population of England in the Eighteenth Century," *J.R.S.S.*, 76, 1912, 261–296. See pp. 279–280.

[5] Henry S. Shryock, Jr., "Methods of Estimating Postcensal Populations," *American Journal of Public Health*, 28, 1938, 1042–7.

[6] Moheau, "Recherches et considérations sur la population de la France, 1778" (republished by Libraire Paul Geuthner, Paris, 1912); Harald Westergaard, *Contributions to the History of Statistics*, London, P. S. King & Son, 1932, pp. 79–99; John Koren, *History of Statistics*, N. Y., Macmillan, 1918, 257–261; Paul E. Vincent, "French Demography in the Eighteenth Century," *Population Studies*, 1, 1947, 44–71.

[7] "The Crop and Livestock Reporting Service of the United States," U. S. Department of Agriculture, *Miscellaneous Publication 171*, 1933; John Koren, *op. cit.*, pp. 642–652.

[8] It can, perhaps, be traced best by reviewing the successive volumes of the publications of the International Statistical Institute and the various national statistical societies. See also, John Koren, *op. cit.*

[9] Faith M. Williams and Carle C. Zimmerman, "Studies of Family Living in the United States and Other Countries," U. S. Department of Agriculture, *Miscellaneous Bulletin, 224*, 1935.

[10] Leah H. Feder, *Unemployment Relief in Periods of Depression*, New York, Russell Sage Foundation, 1936, p. 38.

[11] *Ibid.*, pp. 77–78.

[12] *Ibid.*, pp. 81–83.

[13] *Ibid.*, p. 228.

[14] "Unemployment in the United States," U. S. Bureau of Labor Statistics, *Bulletins 172*, 1915, and *195*, 1916.

[15] Frederick E. Croxton, "Unemployment in Columbus, Ohio, 1921 to 1925," U. S. Bureau of Labor Statistics, Bulletin 409, *Employment and Unemployment Series*, 1926.

[16] J. Frederic Dewhurst and Ernest A. Tupper, "Social and Economic Character of Unemployment in Philadelphia, April, 1929," U. S. Bureau of Labor Statistics, *Bulletin 520*, 1930.

[17] "Unemployment in Buffalo, November 1929; November 1930; November 1931," New York State Department of Labor, *Special Bulletins 163, 167, 172*, 1930–32.

[18] "Unemployment in Syracuse, November 1931," New York State Department of Labor, *Special Bulletin 173*, 1932.

[19] "Hagerstown Morbidity Studies," U. S. Public Health Service, *Public Health Reports, 41,* 1926, 2069–88 (Reprint 1113).

[20] Committee on the Costs of Medical Care, *Report 25-6:* "Incidence of Illness . . . Among Family Groups."

[21] Selwyn D. Collins and G. St. J. Perrott, "The Economic Depression and Sickness," *J.A.S.A. Proceedings,* 29, 1934, 47–51.

[22] Claude E. Robinson, *Straw Votes,* New York, Columbia University Press, 1932, pp. 47 f.

[23] E.g., Stuart A. Rice, *Quantitative Methods in Politics,* New York, Knopf, 1928; W. F. Willcox, "An Attempt to Measure Public Opinion about Repealing the Eighteenth Amendment," *J.A.S.A.,* 26, 1931, 243–261.

[24] A. L. Bowley, "Address to the Economic Science and Statistics Section," British Association for the Advancement of Science, *J.R.S.S.* 69, 1906, 540–557; "Student," "The Probable Error of a Mean," *Biometrika* 6, 1908, 1–25; L. H. C. Tippett, "Random Sampling Numbers," *Tracts for Computers No. XV,* Cambridge University Press, 1927.

[25] E. Keller, "The Principles and Practice of Sampling Metallic Metallurgical Materials," U. S. Bureau of Mines, *Bulletin 122,* 1916; James G. Osborne, "Sampling Errors of Systematic and Random Surveys of Cover-type Areas," *J.A.S.A.* 37, 1942, 256–264; F. X. Schumaker, "Statistical Method in Forestry," *Biometrics Bulletin,* American Statistical Association, 1, 1945, 29.

[26] Adolph Jensen, "The Representative Method in Practice," *Bulletin of the International Statistical Institute,* 22, 1928, pt. 1, p. 386 f.

[27] W. B. Bailey, *Modern Social Conditions,* 1906; F. S. Chapin, *Field Work and Social Surveys,* New York, Century, 1920, p. 116.

[28] Adolph Jensen, *op. cit.,* pp. 394–422.

[29] A. L. Bowley, "Working-class Households in Reading," *J.R.S.S.,* 76, 1913, 672–701; A. L. Bowley and A. R. Burnett-Hurst, *Livelihood and Poverty,* London, Bell & Sons, 1915.

[30] A. L. Bowley, "The Application of Sampling to Economic and Sociological Problems," *J.A.S.A.,* 31, 1936, 474; D. Caradog Jones, "Evolution of the Social Survey in England since Booth," *American Journal of Sociology,* 46, 1941, 818–826; D. Caradog Jones, "The Social Survey of Merseyside," *J.R.S.S.,* 94, 1941, 218–250; P. Ford, *Work and Wealth in a Modern Port,* London, George Allen and Unwin, 1934; London School of Economics, *New Survey of London Life and Labour,* London, P. S. King; 9 vol. 1930–35; A. L. Bowley, "New London Survey," *J.R.S.S.,* 92, 1929, 530–547.

[31] John Hilton, "Enquiry by Sample, an Experiment and its Results," *J.R.S.S.,* 87, 1924, 544; John Hilton, "Some Further Enquiries by Sample," *J.R.S.S.,* 91, 1928, 519–530.

[32] *Bulletin of the International Statistical Institute,* 25, 1931, pt. 2, p. 121.

[33] B. F. George, "A Sample Investigation of the 1931 Population Census," *J.R.S.S.,* 99, 1936, 147; Day Monroe, *Chicago Families,* University of Chicago Press, 1932.

[34] Frank A. Ross, "On Generalization from Limited Data," *Social Forces,* 10, 1931, 32–37; Samuel A. Stouffer, "Sociology and Sampling," in L. L. Bernard (ed.), *Fields and Methods of Sociology,* N. Y., Ray Long and R. R. Smith, 1934; Samuel A. Stouffer, "Statistical Induction in Rural Social Research," *Social Forces,* 13, 1935, 505–515; Frederick F. Stephan, "Practical Problems of Sampling Procedure," *American Sociological Review,* 1, 1936, 569–580; Frederick F. Stephan, "Representative Sampling in Large-scale Surveys," *J.A.S.A.,* 34, 1939, 343–352.

[35] Margaret H. Hogg, "Sources of Incomparability and Error in Employment-Unemployment Surveys," *J.A.S.A.,* 25, 1930, 284–294; Margaret H. Hogg, *The Incidence of Work Shortage,* N. Y., Russell Sage Foundation, 1932.

[36] M. B. Givens, "An Experiment in Advisory Service," *J.A.S.A.,* 29, 1934, 394; "Government Statistics," Social Science Research Council, *Bulletin 26,* 1937, p. 132.

[37] David L. Wickens, *Financial Survey of Urban Housing,* U. S. Bureau of Foreign and Domestic Commerce, 1937.

[38] For example T. C. McCormick, *Comparative Study of Rural Relief and Non-relief Households,* Work Progress Administration, Division of Social Research Washington, 1935; P. M. Hauser, *Workers on Relief in the United States,* March 1935, Work Progress Administration, 1939.

[39] Bureau of Labor Statistics, "Study of Consumer Purchases," *Bulletins 637 to 649 and 691;* Bureau of Home Economics, "Family Income and Expenditures," U. S. Dept. Agr. *Misc. Pub. 339, 345, 356, 370, etc.,* Washington, D. C., 1939.

[40] The Special Committee to Formulate a Plan for a Cost of Living Study, 1929, sought to get an adequate sample of each income class and family type to permit comparisons and regression analysis.

[41] *Michigan Census of Population and Unemployment,* State Emergency Welfare Relief Commission, Lansing, 1936.

[42] J. N. Webb and Albert Westefeld, "Labor Mobility and Relief," *Monthly Labor Review*, January 1939, Reprint R866.

[43] See Conference on Research in Income and Wealth (National Bureau of Economic Research 1943), *Studies in Income and Wealth.* Volume 5, part 2, chapter 10.

[44] Nettie Pauline McGill and Ellen Nathalie Matthews, *The Youth of New York City*, N. Y., Macmillan, 1940.

[45] Conference on Research in Income and Wealth, *op. cit.*, ch. 8.

[46] Calvert L. Dedrick and Morris H. Hansen, "Enumerative Check Census," Vol. IV of the Census of Partial Employment, Unemployment, and Occupations, *Final Report on Total and Partial Unemployment*, 1937, Washington, 1938.

[47] Examples are: H. R. Tolley, "Economic Data from the Sampling Point of View," *J.A.S.A.*, Supplement, 24, 1929, 69–72; Holbrook Working and Harold Hotelling, "Applications of the Theory of Error to the Interpretation of Trends," *J.A.S.A.*, Supplement, 24, 1929, 73–89; Bruce D. Mudgett, "The Application of the Theory of Sampling to Successive Observations not Independent of Each Other," *J.A.S.A.*, Supplement, 24, 1929, 108–117; Jacob Perlman, "Some Problems in the Construction of an Index Number of Rents," *J.A.S.A.*, Supplement, 29, 1934, 25–32; Hugh S. Hanna, "Adequacy of the Sample in Budgetary Studies," *J.A.S.A.*, Supplement, 29, 1934, 131–134.

[48] Committee on Government Statistics and Information Services, "Government Statistics," *Social Science Research Council Bulletin 26*, 1973, pp. 51, 82, 85, 95.

[49] *First Annual Report of the Central Statistical Board*, 1935. pp. 24–25; *Third Annual Report of the Central Statistical Board*, 1937, p. 14; C. L. Dedrick, "Progress of the Census," *J.A.S.A.*, 31, 1936, 737; U. S. Bureau of Agricultural Economics, *Proceedings of Conference on Statistical Methods of Sampling Agricultural Data*, July 14–17, 1936, Iowa State College, Ames, Iowa. (Mimeographed)

[50] "Student," "The Distribution of the Means of Samples Which Are not Drawn at Random," *Biometrika*, 7, 1909–10, 210; T. B. Wood and F. J. M. Stratton, "The Interpretation of Experimental Results," *J. Agr. Sci.*, 3, 1910, 415–440; W. B. Mercer and A. D. Hall, "The Experimental Error of Field Trials," *J. Agr. Sci.*, 4, 1911, 107–132.

[51] For example, T. Lyttleton Lyon, "Some Experiments to Estimate Errors in Field Plot Tests," *Proceedings, American Society of Agronomy*, 3, 1911, 89–114; J. W. Day "The Relation of Size, Shape, and Number of Replications of Plots to Probable Error in Field Experimentation," *Journal of the American Society of Agronomy*, 12, 1920, 100–105; T. Eden, "The Experimental Errors of Field Experiments with Tea"; *J. Agr. Sci.*, 21, 1931, 546; and the bibliography in *J. Amer. Soc. Agron.*, 22, 1930, 1056.

[52] A. R. Clapham, "The Estimation of Yield in Cereal Crops by Sampling Methods," *J. Agr. Sci.*, 19, 1929, 214–235; J. Wishart and A. R. Clapham, "A Study in Sampling Technique," *J. Agr. Sci.*, 19, 1929, 600–618.

[53] A. M. Smith and E. G. Prentice, "Investigation on *Heterodera schachtii* in Lancashire and Cheshire," *Annals of Applied Biology*, 16, 1929, 324–339.

[54] R. A. Fisher, "The Arrangement of Field Experiments," *Journal of the Ministry of Agriculture*, 33, 1926, 503–513; R. A. Fisher, *The Design of Experiments*, London, Oliver and Boyd, 1935.

[55] F. Yates and I. Zacopanay, "The Estimation of the Efficiency of Sampling, with Special Reference to Sampling for Yield in Cereal Experiments," *J. Agr. Sci.*, 25, 1935, 545–577.

[56] Harold F. Dodge and Harry G. Romig, *Sampling Inspection Tables*, John Wiley & Sons, 1944, pp. 1–5.

[57] Walter A. Shewhart, *Economic Control of Quality of Manufactured Product*, New York, D. Van Nostrand Co., 1931; T. C. Fry, *Probability and its Engineering Uses*, D. Van Nostrand Co., 1928.

[58] E. S. Pearson, "Problems of Industrial Sampling," *Supplement to the J.R.S.S.*, 1, 1934, 107–151.

[59] Jerzy Neyman, "On the Two Different Aspects of the Representative Method," *J.R.S.S.*, 97, 1934, 558–625.

[60] J. A. Hubback, "Sampling for Rice Yield in Bihar and Orissa," *Sankhyā*, 7, 1945–46, 281 (reprinted from *Bulletin 166*, Agricultural Research Institute, Pusa, 1927).

[61] J. O. Irwin, "Crop Estimation and its Relation to Agricultural Meteorology," *Supplement* to *J.R.S.S.*, 5, 1938, 1–45.

[62] C. F. Sarle, "Reliability and Adequacy of Farm-price Data," U. S. Department of Agriculture *Bulletin 1480*, 1927; C. F. Sarle, "Adequacy and Reliability of Crop-yield Estimates," U. S. Department of Agriculture, *Technical Bulletin 311*, 1932.

[63] *Proceedings of Conference on Statistical Methods of Sampling Agricultural Data.* U. S. Department of Agriculture, 1936 (Mimeographed).

[64] Irvin Holmes, "Research in Sample Farm Census Methodology," Part I, U. S. Department of Agriculture, *Agricultural Marketing Service*, 1939; Raymond J. Jessen, "An Experiment in the

Design of Agricultural Surveys," *Journal of Farm Economics*, 21, 1939, 856–863; Arnold J. King and Glenn D. Simpson, "New Developments in Agricultural Sampling," *J. Farm Econ.*, 22, 1940, 341–349; C. F. Sarle, "The Possibilities and Limitations of Objective Sampling in Strengthening Agricultural Statistics," *Econometrica*, 8, 1940, 45–62; Arnold J. King and Emil H. Jebe, "An Experiment in Preharvest Sampling of Wheat Fields," *Iowa State College Research Bulletin 273*, 1940; Arnold J. King, Dale E. McCarty, and Miles McPeek, "An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat," *U. S. Dept. of Agr. Technical Bulletin 814*, 1942; Raymond J. Jessen, "Statistical Investigation of a Sample Survey for Obtaining Farm Facts," *Iowa State College Research Bulletin 304*, 1942; Raymond J. Jessen and Earl E. Houseman, "Statistical Investigations of Farm Sample Surveys Taken in Iowa, Florida, and California, *Iowa State College Bulletin 329*, 1944.

[65] Consult the reports in *Sankhyā*, including the pertinent parts of the annual reports of the Indian Statistical Institute; and P. C. Mahalanobis, "On Large-Scale Sample Surveys," *Phil. Trans.*, Series B, 231, 1944, 329–451.

[66] Frederick F. Stephan, W. Edwards Deming, and Morris H. Hansen, "The Sampling Procedure of the 1940 Population Census," *J.A S.A.*, 35, 1940, 615–630.

[67] John N. Webb, M. Starr Northrup, and Stanley L. Payne, "Practical Applications of Theoretical Sampling Methods," *J.A.S.A.*, 38, 1943, 69–77; Lester R. Frankel and J. Stevens Stock, "On the Sample Survey of Unemployment," *J.A.S.A.*, 37, 1942, 77–80.

[68] A. Ross Eckler and E. P. Staudt, "Marketing and Sampling Uses of Population and Housing Data," *J.A.S.A.*, 38, 1943, 87–92; Morris H. Hansen and W. Edwards Deming, "On Some Census Aids to Sampling," *J.A.S.A.*, 38, 1943, 353–357; A. N. Watson, "Use of Small Area Census Data in Marketing Analysis," *Journal of Marketing*, 6, 1942, 42–47.

[69] A. J. King and R. J. Jessen, "The Master Sample of Agriculture," *J.A.S.A.*, 40, 1945, 38–57.

[70] Adolf Jensen, "Purposive Selection," *J.R.S.S.*, 91, 1928, 541–547.

[71] A. R. Mangus, "Sampling in the Field of Rural Relief," *J.A.S.A.*, 29, 1934, 410–415.

[72] Claude Robinson, *Straw Votes*, Columbia University Press, 1932; Claude Robinson, "Recent Developments in the Straw-poll Field," *Public Opinion Quarterly*, 1, 1937, No. 3, pp. 45–56 and No. 4, pp. 42–52.

[73] For example, L. S. Crum, "On Analytical Interpretation of Straw Vote Samples," *J.A.S.A.*, 26, 1931, 243–261.

[74] Daniel Starch, "Factors in the Reliability of Sample," *J.A.S.A.*, Supplement 27, 1932, 190–201; American Institute of Public Opinion, *The New Science of Public Opinion Measurement*, New York (pamphlet); "The Fortune Survey", *Fortune*, quarterly; Elmo Roper, "Sampling Public Opinion," *J.A.S.A.*, 35, 1940, 325–334; Archibald M. Crossley, "Straw Votes in 1936," *Public Opinion Quarterly*, 1, 1937, 24–35; D. Katz and H. Cantril, "Public Opinion Polls," *Sociometry*, 1, 1937, 155–179.

[75] Hadley Cantril (ed.), *Gauging Public Opinion*, Princeton University Press, 1944; Theodore H. Brown, Raymond Franzen, and D. E. Robinson, "Design, Size and Validation for Marketing Research," *Journal of Marketing*, 10, 1946, 221–233; Albert B. Blankenship (ed.), *How to Conduct Consumer and Opinion Research*, Harper, 1946; Ralph Cassady, Jr., "Statistical Sampling Techniques and Marketing Research," *Journal of Marketing*, 9, 1945, 317–341.

[76] *Interviewing for NORC*, National Opinion Research Center, Denver, 1946.

[77] H. Box and G. Thomas, "The Wartime Social Survey," *J.R.S.S.*, 107, 1944, 151.

THE USE OF MARK SENSING IN A LARGE SCALE TESTING PROGRAM*

WALTER L. DEEMER, JR.
AAF School of Aviation Medicine, Randolph Field, Texas

A description is given of the use of mark sensing in the Psychological Program of the Army Air Forces. The principles of mark sensing are summarized and the mark sensing cards of the Psychological Program are exhibited. Some general principles are given for use in deciding on when to use mark sensing in a testing program.

IT IS the purpose of this paper to describe the experience of the Psychological Program of the Army Air Forces in the use of mark sensing, so that this experience may be utilized by others who may have problems similar to those we encountered during the war.

There are five sections to the paper: 1) a brief history of the testing program in which mark sensing was used; 2) a statement of the problems which we attempted to solve by mark sensing; 3) a description of mark sensing; 4) the plan that we devised; and 5) an evaluation of the plan with suggestions for future uses of mark sensing in large scale testing programs.

During the early stages of the war, the testing of aircrew candidates for classification as Bombardier, Navigator, or Pilot was accomplished at three centers, one located at Nashville, Tennessee, one at San Antonio, Texas, and one at Santa Ana, California. A central records agency was established at Ft. Worth, Texas, at Headquarters, AAF Training Command. At the testing centers there was administered to each candidate a battery consisting of twelve to sixteen paper and pencil tests, and five or six psychomotor tests. For a list of the tests administered see Appendix A. The paper and pencil tests were administered to groups of 100 to 300 candidates; the psychomotor tests were administered to groups of four candidates, typically on the day following the group tests. Each psychomotor test required fifteen minutes to administer. For purposes of administrative control, and for research it was desired to have recorded on punched cards in Ft. Worth all the scores of every candidate tested and also certain supplementary information, some of it for identification, such as name and army serial number;

---

* This paper was originally presented at the Educational Research Forum, Endicott, N. Y., August 28, 1947. It represents the opinion of the author only and is not to be construed as representing the official opinion of the War Department or of the Army Air Forces.

some of it biographical, such as age, marital status, education and previous flying training; some of it for classification purposes, such as preference for types of training, status, physical qualifications, recommended classification, and final classification.

In brief, then, the problem was to devise the most efficient and accurate means of getting all the above data on punched cards in Ft. Worth. The obvious method, of course, was to have the necessary rosters prepared at the testing centers and to punch cards from these rosters. On the basis of descriptions of mark sensing, however, it was decided that this method might be more efficient than rosters. Efficiency was important because of the large testing load. An average of over 14,000 men per month was tested during the last half of 1942.

Before describing the plan that was devised, it might be well here to summarize briefly how mark sensing works. Figure 1 shows Group Test Mark Sensing Card No. 1. The particular use of this card in the testing program will be described later. It will be described now as an example of mark sensing in general.

The mark sensing portion of this card begins at column 16. Each of the small ovals with numbers in the center is a mark sensing position and covers three ordinary punching columns. There are 22 mark sensing columns on this card. (A total of 27 mark sensing columns is the maximum that can be put on a card.) There are twelve mark sensing positions in each column. The ovals are immediately above the corresponding punch positions. In use, the ovals are marked with a horizontal line by a special pencil. The cards are processed on the 513 Reproducing Punch, to which a mark sensing unit has been attached. The mark sensing brushes consist of sets of three regular brushes. In each set of three brushes the outer brushes are common and the center brush is connected to a mark sensing brush hub on the plugboard. A mark is sensed when it bridges the gap between the center brush and either outer brush in that group. The marks are sensed conductively rather than photo-electrically. It is necessary, however, to amplify the current through an amplifier unit using vacuum tubes as relays.

The unit is completely flexible in that each mark sensing column may be punched in any column on the card. The card in Figure 1, for example, has been laid out for the first two mark sensing columns to be punched in columns 8 and 9.

The speed of mark sensing is the standard reproducer speed of 100 cards per minute. Ten mark sensing columns can be sensed at one time by one mark sensing unit. Completely to process the card in Figure 1 requires three runs through a reproducer with one mark sensing unit.

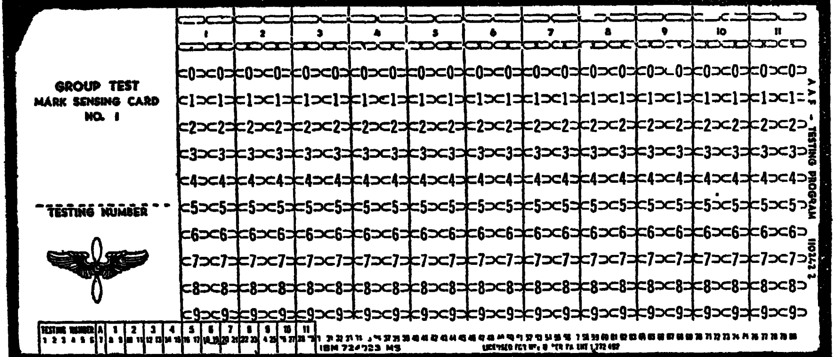


Fig. 1. GROUP TEST MARK SENSING CARD No. 1

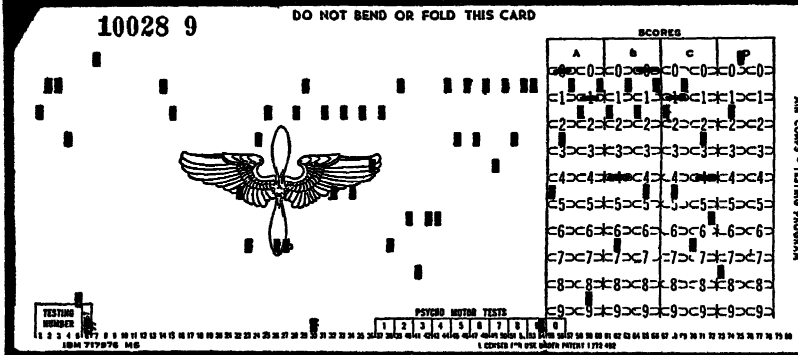


Fig. 2. PSYCHOMOTOR TEST (INDIVIDUAL TEST) MARK SENSING CARD

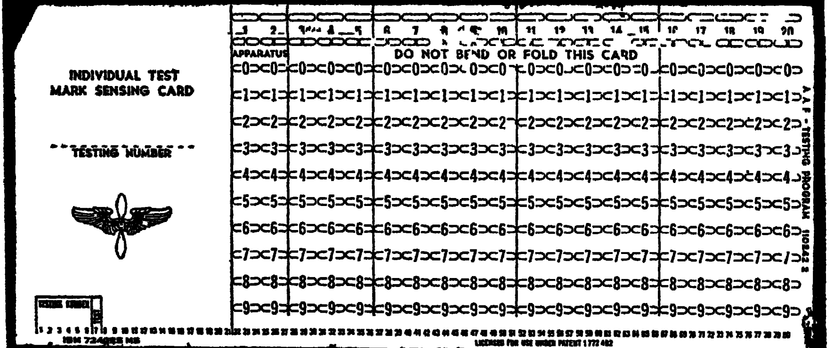


Fig. 3. PSYCHOMOTOR TEST (INDIVIDUAL TEST) MARK SENSING CARD - revised form

In order to verify that a mark has been sensed and punched, use is made of the double punch and blank column detection device, which automatically indicates a card which has a double punched column or a blank column. Since the blank column detection cannot be used without the double punch detection, verification in mark sensing is limited to columns which are designed for single punching only, unless one of the double punches occurs in the 11 or 12 positions, and the other occurs in the 0–9 positions, in which case column splits may be used. The double punch detection device does not detect double *marks* on the cards; it detects double *punching* of the cards.

The original plan for applying mark sensing in The Psychological Program was devised by Colonel John C. Flanagan, who worked out a detailed plan for psychomotor testing. Each candidate was to be given a different card for each psychomotor test. When he entered the room for the test he gave his card for that test to the examiner, who on completion of the test marked the test scores on the card and filed it. The cards for this were planned and printed very early in the program, before it was known definitely just how many psychomotor tests would be used in the battery. Ten cards were prepared for each candidate in order to be sure to have enough. These were card forms like that in Figure 2, which is the number 9 card.

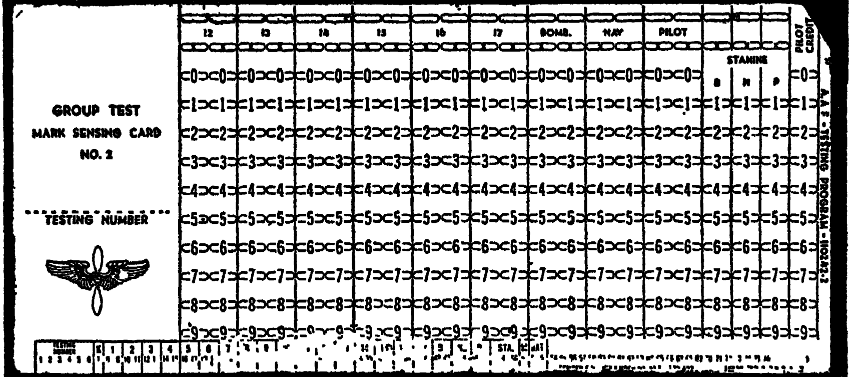
The first five columns of each card contain an identifying number called the *testing number.* In column six is punched the number of the card in the set. The four mark sensing fields with a total of 8 mark sensing columns were for the psychomotor test scores, which for each test consisted of a total raw score, a standard score, and one or two part scores. The four mark sensing fields are in the same position on all ten cards. Each card must therefore be processed by itself, that is, all the number 1 cards were mark sensed and the standard scores for test 1, marked in field A, were punched in columns 37 and 38. For test 2, field A was punched in columns 39 and 40, etc. The marks in the other fields were punched in other columns. After each set of cards was mark sensed all cards were sorted together on columns 1–6 and sequenced on the collator, thus putting all cards of each man together and in order by test. They were then put through the reproducer and by interspersed master gang punching all the scores were combined in the 9-card. The 9-card thus became a summary card. The other cards were stored in semi-dead files, as reference to them was necessary only for verification of scores. All analyses were performed on the summary card. When less than ten psychomotor tests were administered we used less than ten cards, but the 9-card was always the summary card.

Figure 3 shows a revised individual test mark sensing card which allowed more space for recording scores.

It was necessary to make complete plans for reporting scores before any actual scores were recorded in Ft. Worth, and in fact before the extensive testing program was started at the testing centers. After a good deal of discussion with IBM representatives, it was decided that it would be feasible to secure all the test scores and miscellaneous other information by mark sensing. Accordingly two cards were designed for paper and pencil test scores. The first of these is that shown in Figure 1, and the other is Figure 4. There is room on these two cards for 17 two-digit scores. These cards have an advantage over the first psychomotor test card in that the x and y positions are available for marking scores that are unknown, making it possible to insist on a mark in every column, so that the blank column checking device could be used efficiently. Each card could be mark sensed in three runs through the reproducer. Card 1 was reproduced into card 2, making card 2 the summary card. The fields numbered 1 to 17 were for test scores. The three two-digit fields labeled *Bomb.*, *Nav.*, *Pilot,* were for two-digit composite scores giving measures of aptitude for each type of training. These composite scores were based on weighted averages of paper and pencil tests and psychomotor (apparatus) tests. For ease of use, these two-digit scores were converted to single digit scores running from 1 to 9. These single digit scores, called *stanines,* were entered in the three fields near the right edge of card 2. The last column in card 2 had entered in it the amount of credit received by the candidate for previous flying training he may have had. The actual process involved in entering these scores will be described more fully later.

The other two cards used in the testing program are the *name card* and the *stub card.* The name card is shown in Figure 5. The first edition of the stub card is in Figure 6. The second edition is in Figure 7.

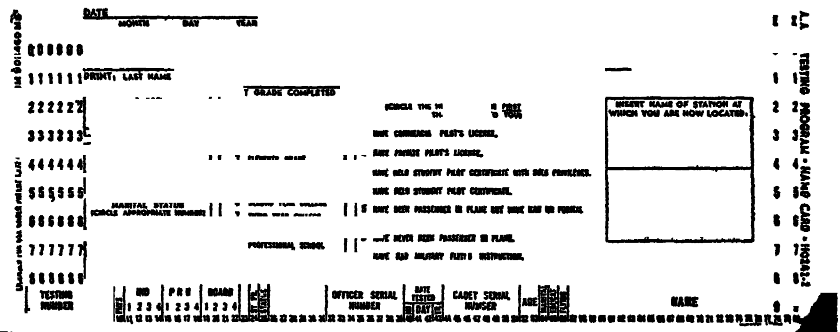
The name card and the stub card were presented to the candidate at the first testing session. They were connected by a small stub and both were pre-numbered and pre-punched with the testing number. Testing numbers were allotted in blocks to the testing centers, and were assigned in chronological order to the cadets, so that a numerical order by testing number was also in order by testing center and by date of testing. The assigning of numbers was automatic since the candidate wrote his name on the name card and the stub and that number was thus assigned to him. His testing number was pre-printed and pre-punched on all his other mark sensing cards, both paper and pencil test cards and psychomotor test cards. A pre-punched, pre-printed card


GROUP TEST
MARK SENSING CARD
NO. 2

TESTING NUMBER

12
13
14
15
16
17
BOMB.
NAV
PILOT
STANINE
B
N
P
PILOT CREDIT
A A F - TESTING PROGRAM

Fig. 4. GROUP TEST MARK SENSING CARD No. 2


DATE
MONTH DAY YEAR

PRINT: LAST NAME

GRADE COMPLETED

MARITAL STATUS (CIRCLE APPROPRIATE NUMBER)

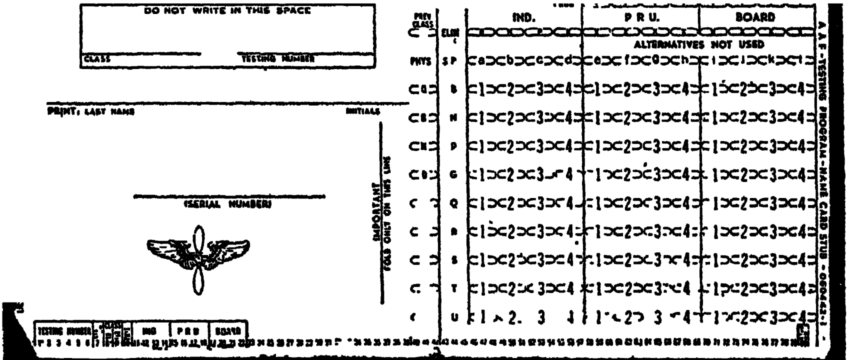
PROFESSIONAL SCHOOL

INSERT NAME OF STATION AT WHICH YOU ARE NOW LOCATED.

HAVE NEVER BEEN PASSENGER IN PLANE.

TESTING NUMBER
IND
P R U
BOARD
OFFICER SERIAL NUMBER
DATE TESTED
DAY
CADET SERIAL NUMBER
AGE
NAME
A.A.F. TESTING PROGRAM - NAME CARD

Fig. 5. NAME CARD


DO NOT WRITE IN THIS SPACE

CLASS
TESTING NUMBER

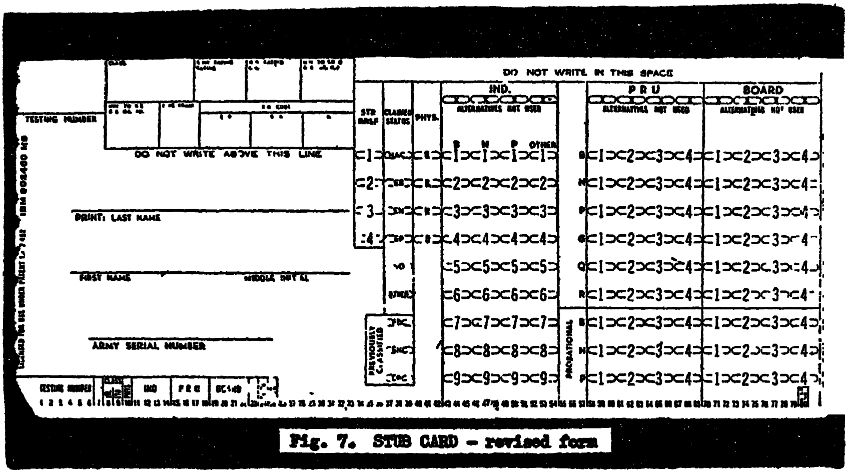
PRINT: LAST NAME
INITIALS

(SERIAL NUMBER)

IMPORTANT
FOLD ONLY ON THIS LINE

PREV. CLASS
ELIM
IND.
P R U.
BOARD
ALTERNATIVES NOT USED
PHYS
SP
B
N
P
G
Q
R
S
T
U
TESTING NUMBER
IND
P R U
BOARD
A A F - TESTING PROGRAM - NAME CARD STUB

Fig. 6. STUB CARD



TESTING NUMBER

DO NOT WRITE ABOVE THIS LINE

PRINT: LAST NAME

FIRST NAME

MIDDLE INIT'AL

ARMY SERIAL NUMBER

DO NOT WRITE IN THIS SPACE

IND.

P R U

BOARD

STN GROUP
CLASSIFIED STATUS
PHYS.
PREVIOUSLY CLASSIFIED
PROBATIONAL

Fig. 7. STUB CARD - revised form

appears in Figure 2. The other cards shown in the figures are samples without the pre-printing and pre-punching.

The name card (Figure 5) did not require any mark sensing. The cadet filled it in by writing the date, his name, army serial number and date of birth, and by encircling the appropriate entries for marital status, education, and previous flying experience. This card was designed for self coding; that is, the punch operator started punching in column 41, with serial number, after date of testing was duplicated in columns 37–40 from a master card. As the card went through the punch the rest of the information became visible.

The stub was torn off by the candidate after filling in name and army serial number, and was retained by him as a record of his testing number, which he wrote on all his group test answer sheets. Prior to the psychomotor test session the candidate turned in his stub and received his set of psychomotor test cards with the same testing number. The mark sensing information on the right of the stub card was filled in at the testing center after the candidate had been tested and classified.

All cards of the candidate were sent to Ft. Worth as they were completed. The name cards were a source of information on the number of people who had started testing. The stub card information was sometimes not available for some weeks or even months after the candidate was tested. The name cards for such candidates were kept in a special file and permitted a rapid count of incomplete processing.

The mark sensing part of the stub card was completed at the testing center as the necessary information became available. The first edition of this card (Figure 6) was quite complex and illustrates a great many of the pitfalls to be avoided in mark sensing. The first mark sensing field gave physical qualification. The *Prev class* space at the top of the first field was marked if the candidate had previously been classified elsewhere. The column of letters beginning with B to the right of the first field gave the labels to be attached to the rows of positions all the way to the right hand margin: B for Bombardier, N for Navigator, P for Pilot, G for Ground, and the other letters for possibilities which might arise. Actually there did arise other classifications, such as flight engineer and gunners. The four columns in the field headed *IND* represented the first to fourth preference of the individual for each type of training. For example, an individual whose first choice was Pilot, second Navigator, third Bombardier and fourth Ground would have had his card marked opposite the P in the first column, opposite the N in the second column, opposite the B in the third column, and opposite the G in the fourth column. The rows are the aircrew positions and the columns are the choices from 1 to 4. In the next field, headed *P.R.U.*, were given the four recommendations of the testing center (Psychological Research Unit); the final classification by the Classification Board was entered in the field headed *Board*.

The top three rows contained special codes not related to those in rows 1 to 9. The row headed *SP* showed the strength of preference for the individual preferences. That is, it indicated how much importance the candidate wanted attached to his stated preferences.

The line above that, headed *Elim*, represented tvpes of training from which the candidate may have already been eliminated. Many men who had been eliminated from Pilot training, for example, were tested for possible entry into Bombardier or Navigator training. The last nine columns in this row were not used. The Y row headed *Prob* represented any training which was recommended or assigned probationally.

It will be seen that double punching on this card occurred in the four columns headed *IND*, and triple punching in this field was possible. Thus if the machine were wired normally, there would be a stop on every card since at least one of the four columns would be double punched. The split column device could not be used to separate the punches because it splits between 11 and 0, and the double punch found here consisted of 0 and a punch from 1-9. Consequently, the double punch detection had to be omitted from this field. Since the double punch blank column detection device on the reproducer will not operate for blank columns without operating for double punches it was not possible to use either of these checks. This turned out to be a very serious difficulty.

Mark sensing is subject to two sources of error, one in the marking and one in the processing in the reproducer. In the stub card we were getting at least our share of errors from both sources.

The directions for marking the stub card were prepared very carefully, but even after some months of use we occasionally found errors in marking due to failure to follow directions; this is not surprising in view of the complexity of the card layout.

By far the largest number of errors in marking the card were due to carelessness, however, rather than to misunderstanding. We found almost every conceivable type of error: columns not marked at all, too many marks, impossible entries, fields transposed.

We found, too, that even correctly marked cards fairly often came out with missing punches or too many punches.

Because we could not use the blank column double punch detection device on the reproducer we had to find the errors by tedious methods, such as running lists on the tabulator for visual checks. This was far from satisfactory either for accuracy or speed.

We also experienced difficulty in processing many stub cards because of their bad condition due to handling. It will be remembered that the candidate kept the card for one day. We had prepared a crease for folding the card so that it could be carried easily in a pocket. The creasing alone did not cause any great processing trouble. But the combination of creasing and handling and carrying in pockets by candidates often did mutilate the card so that it was difficult to process. We had further processing difficulties because of the serrated edge when the card had been detached from the stub. Unless constantly watched and checked the punches from the mark sensing would be slightly out of position. This would eventually cause trouble in the reproducer and collator.

After a few months of use we revised the stub card. The revised card is shown in Figure 7. In the new edition all double punching and all blank columns have been eliminated. That is, each column can be marked correctly in only one place, and for every candidate there will be information for every column.

One reason for revising the stub card was an administrative change in reporting preferences. The candidate was asked to show his strength of interest in each aircrew position on a one to nine scale, instead of ranking his preferences from one to four. Nine positions were therefore necessary for each category (B, N, P, other), making necessary a change in the card dimensions used for the positions. That is, where in the first card we had B, N, P as rows, we now have them as columns.

We could not change the dimensions of PRU and Board Classifications because of the extra data which we had to insert. Hence we were left with three fields—individual preference, PRU recommendation, and Board Classification—all of which should logically have aircrew position as either rows or columns, but which actually had different dimensions. This caused us some difficulty in preparing rosters for outside use.

The cards on which the psychomotor test scores and the paper and pencil test scores were marked were much simpler than the stub card and caused less trouble in processing. Nor was there any apparent misunderstanding of directions for marking. Certain difficulties were encountered with these cards, however. At first the accuracy of entering scores on mark sensing cards was a good deal less than the accuracy of entering scores on a roster. This was discovered by having two crews alternate on both methods of entering. The accuracy was improved by careful supervision, but all testing centers reported a very definite dislike on the part of the clerks for entering scores on mark sensing cards and this probably contributed to inaccuracy.

It should be noted that the personnel who were doing the recording were well trained and highly competent. As a matter of fact, many of the clerks were enlisted men who were probably much more intelligent (and more interested in their work) than the usual clerks who would ordinarily be used for this work. Perhaps the high level of intelligence (as shown in part by high Army General Classification Test scores) contributed to the dissatisfaction felt with the tedious process of marking the cards.

We encountered some informal objection to mark sensing from one or two senior officers who stated that a permanent original record of scores should not consist of marked cards but should consist of rosters. In the case of paper and pencil tests, this objection was unfounded, since the original record was the IBM answer sheet. In the case of the psychomotor tests, however, there was some merit in the objection, since according to the plan for use of the mark sensing cards no other record of the score would be kept. This objection was never communicated formally.

This objection to the use of mark sensing cards in psychomotor testing was reinforced by the need for keeping records for research purposes of many part scores on the psychomotor tests. There was no room on the mark sensing card for these scores, and furthermore the testing unit wanted a record of the part scores for its own use. After a short while, therefore, we found that what could have been a real labor saving method had it been usable as planned, was becoming a burden because it was being superimposed on the old method; that is, score sheets for each psychomotor test were being prepared at the time of the test and the mark sensing cards were being filled in later.

This duplication was soon evident in all the mark sensing operations. For local administrative purposes rosters of the information on the stub card were needed sooner than they could be furnished from Ft. Worth. The result was that the examining centers were making rosters of the same information they were putting on the mark sensing cards. Since the centers were having difficulty in securing sufficient clerical help, the extra labor of making mark sensing cards was a distinct burden.

Certain difficulties were encountered with mark sensing in Ft. Worth. One difficulty was in storing the cards which were not used as summary cards. These consisted of one paper and pencil test card and nine psychomotor test cards, a total of ten cards per man. These cards were the original records which had to be referred to in case the scores on a summary card were questioned. They therefore not only had to be preserved, but had to be readily available because many questionable scores were found on summary cards as they were processed. There was thus a large job of filing these cards. Furthermore, reference to these cards was tedious because of the bulk. Finding the score on a certain test for a certain individual was a rather long operation after thousands of men had been tested.

The problem of original records had further ramifications. When a card was mispunched as it was mark sensed, either because of inaccurate marking or machine malfunction, a correct card was made. But the mispunched card had to be saved since it was an original record. There developed therefore an extensive file of "spoiled originals." When a search of the main file showed a remade card, it was necessary to search the file of spoiled originals for the original card.

Eventually, after a very thorough tryout, mark sensing was abandoned, principally because of the difficulties at the testing units which have been outlined.

The testing program which we were operating was dynamic. The test battery was changed two or three times a year, and other changes such as method of reporting preferences had to be made. The kinds of supplementary information which was needed changed from time to time. We found that new mark sensing cards incorporating the changes took a good deal of time to print. It was difficult to anticipate changes in the testing battery in time to have new cards prepared.

A summary of the factors which we found to be important in the use of mark sensing is given below:

First, and probably most important from the operational side, is complexity of layout. If a complex layout is unavoidable it should be a warning signal against mark sensing. Complexity of layout may consist either in complex codes which lead to difficulty in the marking, or in double punches and blank columns which lead to difficulty in the processing.

Secondly, mark sensing is contra-indicated when a permanent original record is necessary. There are two reasons for this: first, the cards are more bulky to file than typed rosters, and, second, the cards are hard to establish as original records in cases where scores must be verified for agencies outside the testing program.

A third point to consider is whether a record in other form than a mark sensing card is needed before the cards can be processed and a roster prepared by machine. If such a record is needed it may be more economical to use this record for manual punching.

A fourth point is the time necessary for revising cards in the case of a dynamic program. If changes are going to be at all frequent, a mark sensing system may be a liability, since a change can be effected no faster than new cards can be designed and printed.

I want to emphasize and make explicit a point that has been implicit throughout the preceding discussion, namely, that our final decision to give up mark sensing is not a reflection on the value of mark sensing when it is used under conditions for which it is suitable and for which it has been designed. I have felt only that by presenting our experiences it would be possible to delineate more exactly some conditions for which mark sensing is not suitable.

APPENDIX A

Tests used in the Aviation Psychology Program of the AAF for the selection and classification of Aviation Cadets.

For a complete description of the tests and of the testing program see the Research Reports of the Army Air Forces Psychological Program, U. S. Government Printing Office, 1947.

Paper and Pencil Tests

1. Technical Vocabulary
2. Speed of Identification
3. Mechanical Principles

4. Mechanical Information
5. Mathematics
6. Arithmetic Reasoning
7. Numerical Operations
8. Spatial Orientation
    a. Matching aerial photographs with other aerial photographs
    b. Matching aerial photographs with maps
9. Reading Comprehension
10. Dial and Table Reading

Psychomotor Tests

1. Complex Coordination
2. Discrimination Reaction Time
3. Two Hand Coordination
4. Rotary Pursuit
5. Aiming Stress
6. Finger Dexterity
7. Rudder Control

A PROPOSED BASIC COURSE IN STATISTICS

GEORGE W. SNEDECOR
Professor of Statistics, Iowa State College

INTRODUCTION

A BASIC introductory course in statistics was advocated by the National Research Council Committee on Applied Mathematical Statistics in its stimulating report of May, 1947, Reprint and Circular Series, Number 128. The functions of such a course are stated as follows: "First, it should form part of a general education and as such it should be self-contained. Secondly, it should provide essential training for students majoring in the natural and social sciences, which may be developed further in later courses. Finally, it should interest promising young students in statistics as a profession."

Concerning laboratory work in statistics, the Report contains these penetrating remarks: "Often the main objective of such laboratories is the calculation of means, variances, correlation coefficients and other statistical quantities from numerical data of various types. The laboratory work should place more emphasis on interpreting or drawing inferences from data and on the nature of those inferences. Simple experiments should be devised for illustrating probability laws of various kinds, for carrying out sampling operations and other random processes. The traditional flipping of coins and rolling of dice are not adequate for illustrating many important random processes. The mathematical theory of many of these processes is too complicated to be handled at an elementary level; their experimental demonstration will give the beginning student some feeling for their significance."

The writer finds himself in entire agreement with this report of the Committee on Applied Mathematical Statistics. He finds, also, that many other teachers realize the need for revision in the traditional methods of presenting the subject. These statisticians are not only striving to reorient their own ideas but are experimenting with new methods of teaching. It seems timely to propose a concrete program which may serve as a springboard for discussions.

Two merging trends emphasize the necessity for changes in the teaching of statistics. One trend is the astonishing growth in the impact of statistics on society. For many years after the American Statistical Association was founded in 1839, popular knowledge of the subject was confined almost entirely to governmental statistics and the data of economics, though insurance was beginning to enter its present dominating position in our social structure. Today statistics has captured the popular fancy. The extensive data on the sports pages, the various opinion polls, popularity tests of many kinds including those of radio programs, the cost-of-living indexes with their repercussions on wage policies, surveys of consumer preference, crop estimates, quality control—these are some outstanding examples of the preoccupation of people with statistical evidence.

The second trend merging with the first is the no less astonishing growth of statistical theory. Estimation and the testing of hypotheses have clothed with living flesh the dry bones of numerical data. The emphasis on sample-to-population inferences has put new meaning into statistical terminology. The Fisherian concept of "information" has shifted interest from the formal rules of calculation and summarization to the vital processes of getting information into the data by use of appropriate sampling and experimental designs.

The merging of these two trends into a single stream makes new demands on our statistical personnel. Neither the isolated theorist nor the submerged practitioner is able to keep abreast of the current. The specialist in mathematical statistics must acquaint himself with practical problems, and equally, the work-a-day statistician must familiarize himself with the logic of statistical theory: fortunately, this is being made more readily available to those who do not have the benefit of mathematical symbolism. It is only by the intermixture of the two streams that statistics can freely flow onward.

The fact should not be overlooked that the recent upsurge of interest in statistics is based on an ancient pattern of thinking, a form that has developed along with the thought process itself and that is more primitive and more extensive than the logical forms. Concepts of type and of departure from type are counterparts of statistical averages of location (means, regression coefficients, etc.) and of measures of scale such as the interquartile interval and standard deviation. The process of sampling is deeply embedded in human actions. Judgments about probability together with consequent behavior determine much of the pattern of our daily lives. Anyone entering the profession of statistics, especially the teaching of it, should enjoy the confidence that he is engaging in a fundamental activity of mankind.

Since statistics is so intimate a part of our social organization, the conclusion is inevitable that it should be taught generally to our young people. As indicated in the National Research Council Committee report, the elements will doubtless be introduced into the high school curriculum as soon as there is a sufficient supply of trained teachers. Meanwhile, the basic course, taught at the freshman or sophomore level, should serve as the foundation of college curricula to produce such teachers along with other professional statisticians.

Essentially, my idea is to bring the student into awareness of and harmony with the statistical content of our society. This content is extensive. It includes gossip, news and probability theory; sports and old age pensions; gambling and weather prediction; birth rates and living costs; the stock market and epidemics. Such large-scale activities as insurance and the census must be integrated with the more academic concepts of probability, distributions, sampling, estimates and tests of hypotheses.

In working out the following syllabus, the author has set himself the ideal of presenting sound statistics in an interesting fashion. It would seem inadvisable to introduce so vital a subject in an austere and forbidding style. On the other hand, one must be constantly on guard lest he produce false impressions that will later have to be tediously eradicated. Only the elements of statistical thinking should be incorporated in this basic course, but the elements should constitute strategically chosen timbers that will fit readily into the projected structure.

It is a deep-seated conviction of the author that the tools of statistics should be brought out and sharpened only after the need for them has been felt. Tables, graphs, and calculations of the various averages have all too often been presented as the subject matter of statistics rather than as its implements. The very young people whom we wish to enlist, those with imagination and intellectual enthusiasms, may be alienated by dull routine if introduced before any necessity for it becomes evident.

It is clear that no more than an outline of a basic course can be suggested now. The outline will be changed and filled in by the efforts of many teachers who will gain experience from trying the experiment. Some have already started; the important thing now is to get more people working at the job.

PROPOSED OUTLINE OF COURSE

I. *Introduction.* The interest people show in counting and measuring. Some historical items showing the antiquity of the habit. Statistical form of much human thinking. Uncritical attitude of people toward numerical statements. Illustrate fallacies. Interesting and uninteresting statistics. Illustrative material: batting averages, birth rates, election returns, opinion polls, vital statistics, the average man.

II. *Inquiry by Sampling.* Propose a contract to find out for a radio station the number of listeners to Program X. Develop limitations in extent of inquiry and in number of units interviewed. Discuss design of sample, including size. Quota sampling, random sampling, area methods. Use of mail and telephone.

Assume sampling completed and ballots counted. Contrast known fraction of listeners in sample with unknown fraction in the population. Take confidence interval from table and explain meaning. Expand to population total number of listeners, using census information about area sampled.

This sampling problem has been introduced ahead of the chapter on a priori probability for three reasons: (i) It is the newer and more practical concept; (ii) It leads immediately into a modern social problem instead of into the more ancient and perhaps less honorable one of games of chance; (iii) The term *population* has its obvious and fundamental meaning, sample-to-population inferences being inevitable. Experience has convinced me that the student feels himself on familiar ground.

Laboratory: Conduct opinion poll sampling of student body; construct questionnaire, design sample, discuss interview technique, tabulate results, make inferences about population. Students enjoy this experience, ask many penetrating questions and open opportunities for sound instruction.

III. *Random Sampling from Population with Known Constitution.* Use bowl containing equal numbers of beads of two colors, or use table of random numbers with equal numbers of odd and even digits. Draw many samples of 10, recording numbers of "successes." Contrast sample ratios with that of population. Set confidence interval from each sample and determine proportion of correct statements made. Compare results with theory. Emphasize sampling variation and observe improved reliability when samples are combined into larger samples. Compare with poll of radio listeners where parameter is unknown.

Laboratory: Extend sampling experience by use of coins or dice. Record result of each toss for later use.

IV. *Frequency Distribution.* Tabulate numbers of samples with 0 to 10 successes. Present distribution graphically. The mode as an estimate. The small fraction of extreme values as an explanation of confidence in sampling. The mean number of successes in the aggregate of each student's sampling. Emphasize the greater reliability and utility of the mean as compared with the mode. Combine samples of 10 into larger groups and observe (i) the greater variation in the number of successes and (ii) the lesser variation in the proportion of successes. Law of Large Numbers.

V. *Insurance.* Next to government, the greatest cooperative social enterprise in America. Historical items. The premium, the expectation of loss, is the price of protection equally enjoyed by all participants. Life insurance and mortality tables. Expectation of life. Emphasize definite statements about uncertain events. Pure premium for single year term insurance. Level premiums for term and life policies. Insurance vs. savings accounts.

Laboratory: Calculating machines will be required to compute various premiums.

Insurance is now a social problem of great importance. The soundness of the trend away from pure insurance which is predominantly statistical, towards various forms of savings devices which are mainly financial, is questionable. The student should be able clearly to distinguish between insurance and investment.

VI. *A Priori Probability I: Games of Chance.* A primitive, universal human interest. Probability and expectation. Conditions of fair play. Effects of limits on resources and time—sample vs. population. Playing against The House. Systems of play. Probabilities of runs.

Laboratory: Try a system like the Martingale, noting winners and losers at end of specified number of throws. Balance accounts in the aggregate. Work out probabilities in some game.

This is one of the chapters that affords great opportunity for exposing popular fallacies.

VII. *A Priori Probability II: The Binomial Distribution.* Develop first with probability of ½. Compare sample distributions with theoretical. Raise question of testing hypotheses. Develop distribution with probability different from ½. Skewed distributions. Mean and mode.

Laboratory: Throw dice to get data for samples for asymmetrical binomial. Compare samples with population. Save data for testing in next chapter.

VIII. *Empirical Probability.* Prevalence of vague judgments and actions based on them. Sampling basis of most calculated probability. Sampling from populations with unknown probability. Illustrate by throwing a loaded die. Calculate chi-square for hypothesis of perfect balance. Get distribution of chi-square from samples of 10. Test of hypothesis. Meaning and use of table of chi-square. Contrast samplings from populations with known and unknown probability. Emphasize the more practical problem of sampling from unknown probability with resulting inferences and tests. Ordinarily the parameter is forever unknown, but pertinent hypotheses intrude themselves.

Laboratory: Extend the experimental basis for the chi-square distribution. Test an ample number of hypotheses about various samplings that have been made. Extend to more than one degree of freedom. Test the goodness of fit of samples from binomial distributions.

This chapter is the climax of the first part of the course. A substantial body of statistical theory has been accumulated together with some practical problems and a number of socially advantageous applications. Sampling from specified populations has been emphasized. The binomial distribution has been developed. A sampling distribution of chi-square has been built up, leading to confidence in the table. Estimates and tests of hypotheses have been justified and applied. Uncertain inference has been exemplified. If the student ends his contact with the course here he will have had experience with the fundamental concepts of statistics.

IX. *Measurement.* Start with guesses at the length of an 18–20 inch bar. Distribution of guessed lengths. Measure with scale and summarize in distribution. Emphasize variation as characteristic of all measurement. Parameter is again always unknown. Develop idea of normal distribution as a model like the binomial. Properties of the normal distribution. Mean and standard deviation of sample as appropriate estimates of parameters. Methods of calculation. Repeated measurements of the same thing compared to measurements of the members of a sample from a normal population. Emphasize conceptual character of models and impossibility of learning parameters by sampling.

Laboratory: Let all measure height of one member of class. Distribution of heights of men and women. Perform some psychological experiment leading to near-normal distribution.

X. *Sampling Distributions.* Each member of class draws 10 or more samples of 10 from a normal population, using table of random digits for randomization. Mean and variance are calculated for each sample. Tabulate distribution of each, showing normality of first and skewness of second. Estimate population mean and variance from each. Calculate t for each sample, using known population mean. Distribution of t. Calculate confidence interval based on each sample and verify fraction of correct statements.

Laboratory: Construct normal curve from parameters of sampled population. Fit normal distribution to distribution of means. Test normality of distribution of variance.

XI. *Sampling From Some Human Population.* If time and facilities are available, this may be an actual sampling. Usually this chapter will be limited to tabulation, summarization and presentation of available data from some sample survey. Calculate mean and variance. Test normality of distribution. Confidence intervals. Expand estimates to population totals.

Laboratory: Most of this chapter is the laboratory type of work.

XII. *Non-Normal Distributions.* Rectangular and skewed. The median as an estimate. Sample distributions of mean and median. Emphasize that no known actual distribution is normal. Use of median and related order statistics.

Laboratory: Draw complete set of samples from some small rectangular populations showing central tendency of means.

XIII. *Regression.* Growth curves and economic trends. Calculation of linear regression on several assumptions. Deviations of individuals from trend; case studies. Estimates, confidence statements and tests of hypotheses.

Laboratory: Get results of aptitude tests and college grades. Estimate latter from former. Each student calculates his own deviation.

XIV. *Index Numbers.* Construct a simple cost-of-living index with direct economic meaning—the changing cost of a specific bill of goods. Emphasize necessity of examining computation before attaching meaning to an index.

Laboratory: Construct cost-of-student-living index in your college. Continue from year to year.

XV. *Correlation* in normal bivariate population. Estimates. Confidence interval and test of hypothesis $\rho=0$. Mental tests. Inherited characteristics. The variance of differences. Correlation and regression. Rank correlation in samples from non-normal populations.

Laboratory: Construct sampling distribution of r in small samples.

XVI. *Sampling from More than One Population;* e.g., men's heights and women's. Combining estimates—under what circumstances is it appropriate? Stratified sampling. Estimates of mean and variance—weights. Analysis of variance in groups.

Laboratory: Draw samples and verify various estimates. Distribution of F.

The foregoing chapters constitute a course occupying two quarters. With omission of chapters V, XI, XIV, and the latter parts of some other chapters, a semester's course is available. For a third quarter, one or more of the remaining topics may be expanded to suit special groups; quality control for engineers and industrial economists, assays for entomologists, etc.

XVII. *Statistical Instruments of the Federal and State Governments.* The census with population studies. Public health and vital statistics. Marketing and other economic statistics. Crop estimates. Relief and security statistics.

XVIII. *Financial and Business Statistics.* Markets and market records. Trends and forecasts. Consumer acceptance and preferences.

XIX. *Quality Control.* Historical items. Control chart with statistical features. Contracts involving control contrasted with those providing inspection. Sequential test.

XX. *Assays.* Governmental regulations. Pure food and drugs. Insecticides. Vitamins. Purity of seed.

XXI. *Experimentation.* Science and experiments. Alternation of induction and experiment. Control of extraneous variation. Groups and "randomized blocks." Analysis of variance. Statistical control—covariance. Broadening the basis of inference—factorial design. Interaction.

XXII. *The Statistical Attitude.* Group vs. individual. The usual vs. the exceptional. News vs. everyday life. Detection of fallacies. Probability vs. certainty. Evidence vs. proof.

The mathematical accompaniment of such a course can be adapted to the training of the students. It may vary from simple algebraic derivations of formulas with practice in the notation to a full course in mathematical statistics. Since the logical concepts of statistical theory can be presented and verified experimentally, without the symbolism of mathematics, it is the author's opinion that the mathematical formulation should follow some more general presentation such as that outlined in this syllabus.

On writers of texts for this course will rest the heavy responsibility of making their appeal directly to the student. At present there are not enough teachers to go around. An instructor's manual containing background material (such, for example, as the data from a consumer's preference survey) would seem to be a necessity, since this kind of source material is not available to many. In fact, such a manual may well be expanded into a book on the teaching of elementary statistics.

ACTUARIAL ESTIMATES FOR PUBLIC SICKNESS INSURANCE PLANS

ABRAHAM M. NIESSEN*
U. S. Railroad Retirement Board

Proponents of social insurance consider sickness insurance a logical extension of the unemployment compensation program. Two state plans and one Federal sickness insurance plan are now in operation, and a number of other states are giving serious consideration to this matter.

This paper explores the possibility of making cost estimates for public sickness insurance plans along established actuarial lines. The use of the actuarial approach may be of importance when long-range estimates are desired.

Appropriate actuarial techniques are presented for a typical public sickness insurance plan. The method is theoretically simple, and does not require the use of statistics which are not ordinarily available. The details would depend on the provisions of the plan and on the purpose of the cost calculations. The general method is sufficiently flexible to be adoptable in various situations.

IN RECENT YEARS a number of states have considered the introduction of sickness insurance plans for workers covered under the unemployment compensation schemes. Two states, namely Rhode Island and California, actually enacted laws to this effect. A similar development took place with respect to railroad workers for whom a nationwide sickness insurance program was established by the amendments enacted in 1946. The methods used in connection with this type of insurance naturally fall in two categories depending upon the purposes at hand. For short term cost estimates, methods commonly employed by statisticians and economists may be advisable. However, for long range estimates actuarial methods would seem preferable.

To the knowledge of the author no actuarial technique has as yet been developed for long range cost estimates in connection with large public sickness insurance schemes. This paper constitutes an attempt to develop new techniques along lines similar to those used in pension fund work. It is believed that the method here presented is in principle simple and that it can be readily applied without recourse to any special statistics that are not usually available. It is of course recognized that at the present time none of the public insurance plans has been in

---

*** The opinions expressed in this paper are those of the author and do not necessarily represent the official views of the Railroad Retirement Board.**

existence long enough to furnish experience upon which reliable actuarial estimates could be made. Nevertheless, an early discussion of the problem may lead to an exchange of ideas which would result in the establishment of a uniform method for the valuation of plans of this type. It is in the spirit of starting off such a discussion that this paper has been written.

GENERAL CONSIDERATIONS

Sickness rates are known to depend on age, race, sex, occupation, geographical location, employment opportunities, and to some extent also on the duration of coverage. Other factors to be considered are variations by calendar year and secular trends. There is considerable difference between making estimates for a "going" system as compared with a new one, since the sickness experience of different programs will generally be affected by the relation of benefits to wages, administrative procedures, and by the existence of other related social insurance programs. From a short range viewpoint most or all of these factors are important. However, when a long range cost estimate is attempted it is practical to consider only the most important factors. It is believed that a sufficiently reliable estimate may be obtained by considering only variations by age, sex, and race. This approach would seem particularly justifiable in the case of a compulsory state-wide plan provided that experience has been available for a period long enough to be considered representative. Even preliminary estimates can be developed by using actuarial methods. In such a case, the basic rates and averages would have to be estimated on the basis of available sickness rates and previous estimates made in connection with the operations of the unemployment compensation program.

The main advantages of the method presented in this paper lie in the fact that the calculations can be performed with a minimum of statistics, and that provision will be made for the effect of reserves. It is recognized, of course, that a sickness insurance plan in which the right to benefits terminates not later than two years after the year of last employment does not have to cope with the problem of deferred benefits. To this extent, the question of reserves is of much smaller significance than in a retirement plan. However, even in public sickness insurance plans reserves are necessary in order to prevent frequent changes in the level of contributions, and to provide a safety margin for contingencies.

The mathematical development presented in this paper generally follows along the lines of actuarial techniques established for the valuation of large self-insured pension funds. Because of the different nature of the problems involved many significant modifications had to be made. In pension fund work it is customary to consider not only the age of the employees but also the duration of their service. This is taken care of by the introduction of actuarial factors (such as withdrawal rates, wage scales etc.) which depend on both age at entry and duration. It is usually convenient to consider the effect of the duration for only a few years which are then said to constitute the select period. All longer durations are then combined and the corresponding calculations proceed by attained age only. The length of the select period is a matter which has to be determined on the basis of preliminary studies. It is not certain that the duration of covered employment will have a considerable effect on the experience of a sickness insurance program. It was felt, however, that a theoretical discussion of actuarial techniques for such programs should indicate the procedures which would have to be followed in case it is decided to make allowances for the effect of the duration of substantial employment. For this reason, it was deemed advisable to introduce a select period of one year. It would obviously be an easy matter to modify the mathematics so as to eliminate the select period altogether, or to introduce a select period of longer duration.

In order to avoid trivial detail the assumption was made that the reader is familiar with tabulating procedures and that he is acquainted with the rudiments of general actuarial theory. For the benefit of the reader who might wish to refresh his memory on some particular topic a few references have been appended at the end of the paper. Reference (e) has been included mainly because the author of this paper is indebted to Mr. Rusam for a few ideas regarding the construction of actuarial functions for sickness insurance calculations.

Since it was desired to present a general outline of the proposed method, the convenient assumption was made that the only breakdowns will be those by age and duration. If other breakdowns, as by race, sex, and occupation, are deemed necessary, the method would have to undergo some slight modifications especially in the part dealing with the level cost calculations.

DESCRIPTION OF THE PLAN

The sickness insurance plan under discussion is assumed to have the following basic characteristics.

1. Participation is compulsory on the part of all workers covered

under the state unemployment compensation program. No other individuals are admitted.

2. The plan is of the monopolistic type, that is, no substitution of private group insurance is permitted.

3. Contributions are in the form of a level percentage of taxable wages without differentiation by age, sex, race, or occupation.

4. No experience rating is anticipated.

5. The plan is to be financed on an actuarial basis including the maintenance of proper reserves.

6. The only requirement for insurability is the earning of a stated minimum of wages in a given calendar year referred to as the "base year." This stated minimum of wages is thereafter in the text referred to as "qualifying wages."

7. Workers insured at the end of a given base year are eligible for benefits in a succeeding one-year period referred to as the "benefit year." Under the Rhode Island plan, for instance, the base year is a calendar year and the benefit year is a 12-month period beginning in April of the next calendar year. Throughout the benefit year the individual benefit rate remains constant depending on the worker's earnings in the base year or in a specified portion thereof.

8. There is an initial waiting period in each benefit year. Benefits are paid on a weekly or daily basis up for a specified maximum duration.

9. The plan pays cash benefits only when there is a loss of wages as a result of sickness. However, the benefit rate depends only on the total creditable wages in the base year. No medical expenses or death benefits are included.

10. The amount of benefit does not depend on the worker's marital or parental status. In other words, there are no family allowances. It is also assumed that the benefit payments have a uniform duration for every qualified worker.

STATISTICS REQUIRED

It is assumed that with respect to each worker the following data will be available.

1. Social security account number.
2. Sex, race, and occupation.
3. Year of birth.
4. Taxable wages in the base year with special provisions made

for showing quarterly earnings if the benefit is based on earnings in a specified quarter rather than in the whole year.

5. Calendar year of experience (base year).
6. Worker's status as of the end of the preceding base year (insured, not insured).
7. Worker's status as of the end of the base year under consideration (apparently insured, not insured because of failure to earn qualifying wages).

The above is the minimum information to be punched in the wage card. In addition, a claim card is to be maintained with respect to each worker who received benefits in the benefit year corresponding to the base year. This card should contain at least items 1, 2, 3, and 5 of the wage card and in addition

(a) The benefit rate, and

(b) The total of benefits received during the benefit year.

It is believed that such information as listed above, although inadequate for a proper analysis of the sickness experience, will be sufficient for an actuarial estimate along the lines developed in the following sections of the paper. It should be noted that in addition to the analysis of the available statistics it will be necessary to prepare a long range estimate of taxable pay rolls by calendar year.

DEFINITIONS AND NOTATIONS

Throughout the remainder of the paper the following definitions and notations will be used:

n	Calendar year corresponding to the base year under consideration.
x	Mean age in the base year, equals n minus year of birth.
E_x^n	Number of workers born in $n-x$ who earned qualifying wages in the two base years $n-1$ and n. For convenience, this group will be referred to as "insured permanent employees."
N_x^n	Number of new entrants in the base year n at mean age x. Denotes workers born in $n-x$ who earned qualifying wages the base year n but not in $n-1$.
$(wE)_x^{n+1}$	Number among the E_x^n who worked in the base year $n+1$ but earned less than the minimum qualifying wages.

$(wN)_x^{n+1}$	Similar to $(wE)_x^{n+1}$ except that it refers to the N_x^n workers.
$(rE)_x^{n+1}$	Number among the E_x^n who had no wages in the base year $n+1$.
$(rN)_x^{n+1}$	Number among the N_x^n who had no wages in the base year $n+1$.
$(SE)_x^n$	Total taxable wages in n earned by the E_x^n workers.
$(SN)_x^n$	Total taxable wages in n earned by the N_x^n workers.
$(SO)^n$	Total of wages in n earned by all workers who failed to earn in that year qualifying wages.
P^n	Total taxable pay roll in n.
$(BE)_x^n$	Total sickness benefits paid during the benefit year corresponding to the base year n with respect to the E_x^n workers.
$(BN)_x^n$	Total sickness benefits paid during the benefit year corresponding to the base year n with respect to the N_x^n workers.

COMPUTATION OF RATES AND AVERAGES

It is assumed that the statistics described above are available with respect to a period of several base years, say, from c to d. The symbol $\sum\limits_n$ denotes summation from $n=c$ to $n=d$. Benefit data obviously need be available for the benefit years corresponding to these base years, which will thus extend the period for which statistics must be available for a year and a fraction beyond the calendar year d, while at the same time the benefit data for the first year and a part of the next year of the period will be ignored. The total probabilitics of loss of insured status, $\overset{w}{q}_x$, for new entrants and permanent employees respectively are:

$$\overset{w}{q}_{[x]} = \frac{\sum\limits_n \left[(wN)_x^{n+1} + (rN)_x^{n+1}\right]}{\sum\limits_n N_x^n} \tag{1}$$

$$\overset{w}{q}_x = \frac{\sum\limits_n \left[(wE)_x^{n+1} + (rE)_x^{n+1}\right]}{\sum E_x^n} \cdot \tag{1a}$$

The average taxable wages, s_x, by age in the base year, again with a select period of one year, are:

$$s_{[x]} = \frac{\sum_n (SN)_x^n}{\sum_n N_x^n} \tag{2}$$

$$s_x = \frac{\sum_n (SE)_x^n}{\sum_n E_x^n}. \tag{2a}$$

The average benefit payments in the benefit year, b_x, per individual age x in the base year who could be eligible for benefits with respect to that base year, are then computed by means of the following formulas:

$$b_{[x]} = \frac{\sum_n (BN)_x^n}{\sum_n N_x^n} \tag{3}$$

$$b_x = \frac{\sum_n (BE)_x^n}{\sum_n E_x^n}. \tag{3a}$$

A set of r_x's, giving the proportionate age distribution of new entrants, is obtained from the formula:

$$\begin{array}{c} \sum N'_x \\ \\ \sum_n \sum_x N_x^n \end{array} \tag{4}$$

Finally, a ratio g, which represents the portion of the total taxable pay roll which gives no rise to any benefit payments, is computed from:

$$g = \frac{\sum_n (SO)^n}{\sum_n P^n}. \tag{5}$$

THE SERVICE TABLE AND ACTUARIAL FUNCTIONS

All functions (1) through (5) can easily be obtained from proper tabulations. The rates and the averages may first be computed for quinquennial age groups. The values by single age can then be obtained by one of the several standard interpolation methods.

At this juncture it becomes necessary to construct a service table. In the case of a pension fund, the service table would take separate account of exits from service by reason of death, withdrawal, disability, and retirement. If we denote the tabular number in service at age x by l^s_x the 'in service survivors' a year later or l^s_{x+1} would be obtained by subtracting from l^s_x all exits between ages x and $x+1$ as computed on this base. As an illustration, let us assume that $l^s_{25} = 100{,}000$ and that the total of the probabilities of exit is .2. Then, the computed number of exits between ages 25 and 26 is 20,000, and $l^s_{26} = 80{,}000$. For the purpose of the development presented in this paper a special definition of exit from service was introduced. Here, an individual is considered to have left covered employment if for any reason whatsoever he failed to earn qualifying wages in a given base year. The combined probabilities of such exit are defined by equations (1) and (1a). Our particular service table is then constructed as follows:

Begin with a selected young age, say, 15.

$$\begin{aligned} \text{Put}\quad l'_{[15]} &= 100{,}000 \\ \text{Then,}\quad l'_{16} &= l'_{[15]}(1 - q^w_{[15]}) \\ l'_{17} &= l'_{16}(1 - q^w_{16}) \\ \\ l'_{x+1} &= l'_x(1 - q^w_x) \end{aligned} \tag{6}$$

The select section of the service table is then computed by the formula:

$$l'_{[x]} = \frac{l'_{x+1}}{1 - q^w_{[x]}}\,. \tag{7}$$

It will be noted that from the way the table is constructed we have the following relationships:

$$\frac{l'_{x+1}}{l'_{[x]}} = \frac{\text{Those among the insured new entrants } N^n_x \text{ who also earned qualifying wages in the base year } n+1}{N^n_x} \tag{8}$$

$$\frac{l'_{x+1}}{l'} = \frac{\text{Those among the insured permanent employees } E^n_x \text{ who also earned qualified wages in the base year } n+1}{E^n}\,. \tag{8a}$$

We now compute the following actuarial functions:

$$D'_{[x]} = v^x l'_{[x]}, \quad \text{and} \quad D'_x = v^x l'_x \tag{9}$$

where $v = 1/1+i$, i being the interest rate used in the calculation.

$${}^bN_x = D'_x b_x + D'_{x+1} b_{x+1} + \cdots + D'_\omega b_\omega \tag{10}$$

where ω denotes the limiting age which in practical situations will not be more than 80.

$${}^bN_{[x]} = D'_{[x]} b_{[x]} + {}^bN_{x+1} \tag{10a}$$

$${}^bA_{[x]} = v^m \frac{{}^bN_{[x]}}{D'_{[x]}} \tag{11}$$

$${}^bA_x = v^m \frac{{}^bN_x}{D'_x} \cdot \tag{11a}$$

In formulas (11) and (11a), m is the period (in years) between the end of the base year and the middle of the corresponding benefit year.

The ${}^bA_x$ represent the average lump-sum present value of all future benefit payments based on wages in the base year just ended and on future base years during which the individual will earn qualifying wages continuously from the present year on.

LEVEL COST CALCULATIONS

Assume that a valuation of the sickness insurance plan is to be made as of the end of the calendar year n. On the valuation date there will be, say, E_x^n insured permanent employees, and N_x^n insured new entrants. In addition, benefits may still be payable for the remainder of the beneyear with respect to the E_x^{n-1} and N_x^{n-1} groups. Let us furthermore assume that the balance in the account on an accrual basis on the valuation date, that is the funds on hand plus any additional income due but not yet received less disbursements obligated but not yet paid, is R^n.

The present value of the liabilities with respect to persons insured on the valuation date may then be computed as follows:

With respect to the E_x^n insured permanent employees

$$L^E = \sum_x E_x^n \cdot {}^bA_x. \tag{12}$$

With respect to the insured new entrants N_x^n

$$L^N = \sum_x N_x^n \cdot {}^bA_{[x]}. \tag{13}$$

With respect to the workers eligible for benefits in the remaining portion of the current year

$$L^0 = v^{\frac{1}{2}h}\cdot h\sum_x \left(E_x^{n-1} b_x + N_x^{n-1} b_{[x]}\right) f \tag{14}$$

where h denotes the fraction of the remaining current benefit year, and f is a modifying factor to allow for the seasonal variation in disbursements throughout the benefit year.

$$\text{Total} \quad L_1 = L^E + L^N + L^0. \tag{15}$$

At this junction it becomes necessary to compute pay roll functions as follows:

Present value of 1 per cent of all future taxable pay rolls

$$P = .01(1+i)^j \sum_{k=1}^{\infty} P^{n+k} v^k. \tag{16}$$

Here j is less than 1 and $(1+i)^j$ is designed to make allowance for the fact that the contributions on the average come in sooner than at the end of the calendar year. Thus, for instance, if contributions with respect to a given calendar quarter are collected at the end of the succeeding quarter, j may be taken as $\frac{1}{8}$. The individual present value ${}^sP_x$ for an insured worker age x of 1 per cent of future creditable wages taken in account only as long as the individual continues to earn qualifying wages in consecutive years beginning from the present time is obtained from the formula:

$${}^sP_x = .01(1+i)^j \frac{\sum_{k=1}^{\omega-x} D'_{x+k}\cdot s_{x+k}}{D'_x}. \tag{17}$$

The present value on the valuation date of 1 per cent of such future wages for all employees insured by reason of having qualifying wages in the last base year thus becomes:

$$P_1 = \sum_x \left(E_x^n + N_x^n\right)\cdot {}^sP_x. \tag{18}$$

The individual present value as of the end of the year of entry for future entrants, otherwise similar to ${}^sP_x$ is:

$${}^sP_{[x]} = .01 \frac{\left(D'_{[x]}\cdot s_{[x]} + D'_{x+1}\cdot {}^sP_{x+1}\right)}{D'_{[x]}}. \tag{19}$$

The theoretical adjustment for interest on the $D'_{[x]} \cdot s_{[x]}$ component is negligible and can be omitted.

For a group of, say, 10,000 future entrants we have the present value of their benefits (following them only as long as they continue to earn qualifying wages in successive years)

$$L' = 10{,}000 \sum_x r_x \, {}^bA_{[x]} \tag{20}$$

and the present value of 1 per cent of their creditable wages during the period of continued qualifying wages is

$$P'_1 = 10{,}000 \sum_x r_x \, {}^sP_{[x]}. \tag{21}$$

Thus, the level cost of benefits to future entrants as a per cent of their pay roll is

$$t = \frac{L'}{P'_1}. \tag{22}$$

The final cost calculations are then as follows:

Item	*Symbol*	*Formula*
A. Balance in the account	R^n	
B. Liabilities with respect to workers with qualifying wages in either the last base-year or in the previous one	L_1	(15)
C. Present value of 1 per cent of future taxable pay rolls	P	(16)
D. Present value of 1 per cent of future pay rolls with respect to workers who will fail to earn qualifying wages	$g \cdot P$	(5)
E. Present value of 1 per cent of future pay rolls for insured employees taken only so long as they continue to earn qualifying wages, considering only workers with qualifying wages in the last base year	P_1	(18)
F. Present value of 1 per cent of pay roll of future entrants so long as they will continue to earn qualifying wages, (C-D-E)	P_2	
G. Normal level percentage cost for future entrants	t	(22)
H. Value of benefits for future entrants (item *G* times item *F*)	L_2	
I. Total net liabilities ($B+H-A$)	L	
J. Net level cost on the valuation date as a per cent of the taxable pay roll	L/P	
K. Loading for administrative expenses and contingencies (per cent of taxable pay roll)	e	
L. Total cost	$L/P+e$	

CONCLUDING REMARKS

1. The formulas presented before are related to an estimate extending into perpetuity. It is apparent that the same type of formulas can be set down for a calculation limited to a given number of years. Such a modification would not involve any special theoretical difficulties.

2. The service table has been developed without the use of death rates. This approach was taken because in practice reliable statistics on deaths in service would not be available. It should be pointed out that for all practical intents and purposes deaths have been taken into consideration since in the development of the probabilities of loss of insurance the expressions on the right hand side of equations (1) include separations through death. In addition, deaths among individuals insured in a given benefit year were also accounted for by using as the numerators of equations (2) the total sickness benefits actually paid. If the existence of applicable mortality statistics could be assumed, a modified service table could easily be constructed, although such a refinement would probably not be warranted.

3. The level cost calculations proceed on the tacit assumption that the ratio g (formula 5) will remain constant throughout the years. Such will undoubtedly not be the case. However, the ratio g is bound to be small so that cyclical or other variations in its value will not have a significant effect on the results of the valuation. The annual wage data published by the Railroad Retirement Board (Compensation and Service of Railroad Employees) give some indication of the magnitude of this ratio. Thus, in 1945, less than 1 per cent of the total creditable compensation went to employees who earned in that year less than $150.

4. The method makes no allowance for future changes in relationships between benefits according to age, individual wage levels, age distribution of new entrants, and benefit provisions. In this respect, the procedure is similar to the one commonly used in pension fund work. Instead of trying to introduce the probable variations directly in the calculations, it would be much simpler to make two supplementary estimates according to so called "high" and "low" assumptions. Since cost estimates of this kind are subject to considerable errors regardless of the mathematical refinements introduced, the supplementary calculations are in any event necessary for the establishment of a proper range for the cost figure.

ACKNOWLEDGMENT

The author is indebted to Mr. Joseph Musher, Actuary of the Railroad Retirement Board, for many valuable suggestions and for his critical review of this paper.

REFERENCES

[a] Construction of service table: *Life Contingencies*, by E. F. Spurgeon, Cambridge University Press, London, 1932 (Chapters I, II, and XXI).

[b] Actuarial functions: *U. S. Life Tables and Actuarial Tables 1939–41*, Bureau of the Census, 1946 (Part IV).

[c] Graduation and interpolation: *Elements of Graduation*, by M. D. Miller, The Actuarial Society of America and the American Institute of Actuaries, 1946.

[d] Valuation of pension funds:

1. "On Staff Pension Funds," by George King, *Journal of the Institute of Actuaries* XXXIX, pp. 129–184.
2. "Actuarial Valuations of the Assets and Liabilities under the Railroad Retirements Acts," *Annual Reports of the Railroad Retirement Board* for the years 1940, 1943, and 1946.

[e] Mathematics of private sickness insurance: "Grundzuge der Mathematik der Privaten Krankheitskostenversicherung," by Friedrich Rusam, *Transactions of the Twelfth International Congress of Actuaries*, Lucerne, 1940, Vol. IV, pp. 147–167.

EARNINGS OF NONFARM EMPLOYEES IN THE U. S., 1890–1946

STANLEY LEBERGOTT*
Montreal, Canada

Estimates of money and real earnings of nonfarm employees are presented for the years 1890 to 1946, these estimates reflecting the loss in earnings which arises from unemployment. Series are likewise given for full time equivalent earnings of nonfarm employees as a group and for each of the major industry groups for 1919–46, while the Bureau of Labor Statistics and the Bureau of the Census have developed estimates of unemployment for the period 1920–46. The sources of the rise in earnings after World War I are analyzed and it is concluded that about three-fourths of the rise in earnings derived from increased pay on a given job, while one-quarter originated in the changing distribution and characteristics of nonfarm employees.

SINCE the monumental endeavors of Paul Douglas first provided us with estimates of wages for the period 1890–1926, a massive development in the basic materials for income estimation has taken place.[1] The National Bureau of Economic Research and the Department of Commerce have between them provided estimates of full time earnings for the period 1919–1946,[2] and the Bureau of Labor Statistics has developed estimates of unemployment for the period 1920–1940.[3] The present study draws upon these materials and those of Douglas to provide estimates of:

Money and real earnings of nonfarm employees, 1890–1946.

Money earnings of employed workers by major industry groups, 1919–1946.

Attention is also given to the chief sources for the rise in earnings during the period following World War I.[4] Section 1 discusses the general

---

* The views expressed in this article are those of the author and do not necessarily reflect those of any institution with which he is connected. The writer wishes to acknowledge the assistance of Sophia Cooper in the preparation of Tables 5–7.

[1] Paul Douglas, *Real Wages in the United States, 1890–1926*, 1930.

[2] Simon Kuznets, *National Income and Its Composition, 1919–1938*, 1941; *Survey of Current Business, National Income Supplement*, July 1947.

[3] Preliminary BLS estimates for 1929–40 appear in the writer's technical memorandum of July 1945. The estimates for 1920–28 are on a comparable basis and will be released at a later date. 1940–1946 data are from Bureau of the Census, *Labor Force, Employment and Unemployment in the United States, 1940–1946*, 1947.

[4] Throughout this paper references are to employees in nonfarm industry. The term earnings refers to wages, salaries and supplements when used in connection with Tables 1 and 2 and the related text, but does not include supplements in other contexts.

trend of earnings and the sources of the rise in earnings; Section 2 is devoted to a discussion of full time earnings by industry; while Section 3 deals with the actual derivation of the estimates.

The limitations of the basic data used in this study have been discussed in the sources from which they are drawn and are by now well known. But by way of caution it may be reiterated that estimates of unemployment, earnings and cost of living in the years prior to 1930 are inevitably imperfect. The estimates presented in this paper partake of these imperfections. Nevertheless such estimates are required to give us a somewhat better perspective on the course of wage earner incomes than can be gotten from a review of data for the years since 1930. It is believed that the present series does give such a perspective, being of the right order of magnitude, moving in the right general direction and being an improvement over existing estimates for this period.

The money earnings of the average nonfarm employee rose by 251 per cent over the half century, 1890–94—1940–44, while real earnings gained by 68 per cent over the same period. As Tables 1 and 2 emphasize, the gain in real earnings over the decades from McKinley to Coolidge was a most moderate one, the real rise beginning in the much maligned years of the new prosperity. The effect of the depression was so comprehensive, so catastrophic, however, that real earnings in 1930–34 were almost identical with those in 1890–94, forty years before. Despite a vast and persistent gain in wage rates—full time equivalent earnings rose 136 per cent in the interval—the fact that almost a fifth

TABLE 1

EARNINGS OF NONFARM EMPLOYEES
1890–1916

	Full Time Equivalent Earnings	Money Earnings	Real Earnings (1910–14 dollars)
1890–94	$ 553	$ 502	$ 636
1895–99	542	468	597
1900–04	590	554	644
1905–09	641	597	647
1910–14	700	649	649
1915–19	978	918	678
1920–24	1,392	1,235	682
1925–29	1,492	1,384	781
1930–34	1,307	950	640
1935–39	1,366	1,103	721
1940–44	1,910	1,763	1,066
1945–46	2,508	2,408	1,261

TABLE 2

MONEY AND REAL EARNINGS OF NONFARM EMPLOYEES
1890–1946

	Full Time Equivalent Earnings	Per cent of Time Lost by Unemployment	Money Earnings	Cost of Living (1910–14 = 100.0)	Real Earnings (1910–14 dollars)
1890	$ 559	6.2	$ 524	81.4	$ 644
1891	560	6.7	522	79.0	661
1892	567	4.6	541	79.0	685
1893	553	10.6	494	78.5	629
1894	524	18.1	429	76.5	561
1895	542	13.3	470	76.5	614
1896	537	16.6	448	78.5	571
1897	537	15.8	452	78.5	576
1898	542	14.8	462	78.5	589
1899	553	8.4	507	80.1	633
1900	563	7.5	521	81.8	637
1901	579	5.5	547	84.0	651
1902	589	4.7	561	85.7	655
1903	611	4.9	581	89.7	648
1904	609	8.0	560	88.5	633
1905	621	4.6	591	88.5	668
1906	635	3.7	612	91.4	670
1907	659	4.8	627	95.7	655
1908	630	14.2	541	92.8	583
1909	658	6.8	613	92.8	661
1910	685	5.0	651	97.2	670
1911	677	7.0	630	99.7	632
1912	706	4.8	672	100.0	672
1913	733	6.0	689	101.0	682
1914	700	14.1	601	102.6	586
1915	729	13.3	632	103.5	611
1916	826	4.2	791	111.2	711
1917	956	3.8	920	130.8	703
1918	1,107	4.3	1,059	153.5	689
1919	1,260	6.9	1,173	176.8	663
1920	1,448	5.0	1,376	204.7	672
1921	1,347	21.0	1,064	182.4	583
1922	1,330	14.4	1,138	171.0	665
1923	1,413	6.1	1,327	174.1	762
1924	1,423	10.6	1,272	174.5	729
1925	1,450	8.5	1,327	179.1	741
1926	1,479	5.9	1,392	180.5	771
1927	1,488	8.4	1,363	177.1	770
1928	1,507	8.6	1,377	175.1	786
1929	1,535	4.9	1,460	174.9	835
1930	1,495	13.4	1,295	170.5	760
1931	1,410	24.1	1,070	155.2	689
1932	1,251	35.2	811	139.4	582
1933	1,170	36.5	743	131.9	563
1934	1,208	31.2	831	136.7	608
1935	1,261	28.8	898	140.1	641
1936	1,326	24.0	1,008	141.5	712
1937	1,413	20.1	1,129	146.7	770
1938	1,399	26.7	1,025	143.9	712
1939	1,433	23.6	1,095	141.9	772

TABLE 2 (Continued)

	Full Time Equivalent Earnings	Per cent of Time Lost by Unemployment	Money Earnings	Cost of Living (1910–14= 100.0)	Real Earnings (1910–14 dollars)
1940	1,470	20.1	1,175	143.1	821
1941	1,623	13.6	1,402	150.2	933
1942	1,896	6.4	1,775	166.4	1,067
1943	2,195	2.6	2,138	176.5	1,211
1944	2,365	1.6	2,327	179.2	1,299
1945	2,456	2.6	2,392	183.4	1,304
1946	2,560	5.3	2,424	198.9	1,219

of our labor force was without work had completely nullified the gain in rates. Within the next ten years the advent first of a qualified prosperity and then of a war production program had changed famine to feast; real earnings rose by 67.0 per cent as compared to the 0.6 per cent gain over the previous four decades. Money earnings doubled in the six years following the outbreak of both World War I and II. But real earnings rose 69.0 per cent in the later period, thanks to effective price control, as compared to a mere 15 per cent in the earlier period.

The scrupulous reader will keep in mind the inadequacies of any cost of living index which, even at best, can only take inadequate account of quality changes. But though doing so he may choose to set off against such errors the higher costs of urban existence which a cost of living index will likewise tend to ignore. While the measurement of family income is a stage removed from our present concern it seems likely that the trend of that series would have been similar to that of the present series.[5] And as the present series are not measures of family income, so they are not measures of labor costs. Information on manufacturing for the 1920–43 period, for example, very definitely emphasises the point that labor costs may decline sharply at the same time that wage rates and annual earnings gain just as sharply.

While it is not possible to establish the chief sources of the rise in earnings over the entire half century a very rough estimate can be made for the years 1919–23 to 1939–43 when money earnings rose some 18 per cent. The over-all rise in full time equivalent earnings had been 21.6 per cent, 4 points of which had been offset by increased unemployment, leaving 17.6 per cent as the rise in actual money earnings. (It is of interest that so sizable an offset for unemployment occurs: the earlier years included the entire postwar depression while the 1939–43 period included at least 3 years of substantial defense and war production.)

---

[5] In 1901 the head of family provided 80 per cent of the family's income: in 1935–36 the figure was 82 per cent. The 1901 estimate is based on data appearing in the Commissioner of Labor's Report (1904), pp. 362, 366. The 1935–36 estimate is based on ratios to be derived from the various BLS Consumer Purchases volumes, these being weighted by the distribution of nonfarm families given in *Consumer Incomes in the United States*, National Resources Committee, Table 24 B.

Returning to the 21.6 per cent rise in full time equivalent earnings, what can we say of the sources for that rise? An exceedingly rough approximation suggests that upwards of three-quarters of the gain came about because of increased earnings on a given job, while the balance derived from changes in the composition of the labor force. Three elements are significant in this composition shift: industrial change, occupational change, and age change.

The importance of the industrial change may be approximated by using a constant 1919–23 industrial pattern for both periods in place of the actual ones. Tables 3 and 4 present employment and earnings estimates for both 1919–23 and 1939–43.[6] By assuming a constant distribution of weights between the manufacturing sub-groups and between the major industrial groups we secure an estimated income gain of 1.6 points less than the actual gain. In other words 1.6 out of the 21.6 points gain in full time earnings may be assumed to have derived from the change in industrial composition.

But more than a shift in industrial composition was at work. At the end of the period the average worker was more highly skilled. What portion of the rise in earnings stemmed from this source? Table 5 presents the data used under the assumption of a constant occupational composition of the employee work force. Occupation data are readily available and as here presented exclude both professional workers and retail trade proprietors, managers and officials, since so large a proportion of both groups is composed of self-employed persons.[7]

Earnings data are less readily come by, and even conceding a number of points we must assume that a comparison can usefully be made by taking 1920 and 1939 occupational earnings as the same. (It can be demonstrated that using such an assumption for the industrial data, where we do have earnings at both dates, makes little difference in the results.) We can fairly well estimate what proportion of the experienced workers in the labor force who were employed 12 months in 1939 were not wage and salary workers—hence ones whose incomes would not be representative of employees earnings in a given occupational group.[8] By assuming these workers to cluster at the lower end of the wage and salary income distribution we can then derive median wage and salary income of full time employees in the various occupational groups.[9] Combining these medians first with a constant and then a varying occupational distribution we arrive at practically identical figures for 1920 and 1940. Taking account of the differences between the 1920 distribution and the 1919–23 distribution, as well as those between 1940 and 1939–43 we can conclude that there was probably a slight, though insubstantial, gain because of occupational change.

---

[6] The employment data are derived from Kuznets and The *National Income Supplement*, with adjustments for comparability as described below in section 3, where a description of the derivation of earnings data likewise appears. For manufacturing, the 1939–43 income using constant sub-group weights was $1,693 as compared to the actual $1,764.

[7] Alba Edwards, *Comparative Occupation Statistics, 1870–1940*, 1943, pp. 186–187. For 1940 the number of employed is used; for 1920, the number of gainfully occupied. It is assumed that the changing distribution of self-employed persons within the specified occupational groups would not materially affect the results.

A third factor to be considered as a source for the rise in earnings is the changing age distribution of the labor force. By comparing the urban workers age distribution of 1920 with that of 1940 and assuming

TABLE 3

EARNINGS AND EMPLOYMENT BY MANUFACTURING INDUSTRY

1919–23 and 1939–43

	Average Full Time Equivalent Employees		Average Full Time Equivalent Incomes	
	1919-23	1939–43	1919–23	1939–43
Food	849	1,299	1,412	1,551
Tobacco	166	102	983	1,140
Textile	1,161	1,306	1,076	1,209
Apparel	618	1,003	1,305	1,226
Lumber	579	551	1,102	1,113
Furniture	163	423	1,263	1,372
Paper	233	360	1,340	1,689
Printing	442	566	1,609	1,893
Chemicals	448	619	1,567	1,948
Petr. coal	111	165	1,844	2,227
Rubber	120	181	1,523	1,901
Leather	316	385	1,283	1,284
Stone	324	399	1,393	1,620
Iron	663	1,709	1,594	2,007
Nonferrous	320	396	1,464	1,951
Machinery	1,514	2,943	1,557	2,193
Auto	362	513	1,674	2,359

---

[8] From the 1940 Census of Population, *Occupational Characteristics*, Table 13, we can compute the per cent of wage or salary workers in each occupational group that worked 12 months. Applying these percentages to the number of wage earners in the complete count (1940 Census of Population, *The Labor Force*, Part I, Table 72), deducting the result from the total of experienced persons in the labor force 12 months, we have the number of non-wage or salary workers that worked 12 months in 1939.

[9] Medians were computed by deducting the number of non-wage or salary workers who worked 12 months from the total experienced labor force who worked 12 months, cumulating from the $0 income group. (Income distribution from *The Labor Force*, *loc. cit.*) The clustering premise is clearly inadequate. Just as clearly its bias is not going to make much difference in the medians.

TABLE 4

EARNINGS AND EMPLOYMENT BY MAJOR INDUSTRY GROUPS
1919–23 and 1939–43

	Average Full Time Equivalent Employees		Average Full Time Equivalent Income	
	1919–23	1939–43	1919–23	1939–43
Mining	1,032	928	$1,696	$1,658
Construction	1,245	1,594	1,482	1,787
Trade	4,074	6,587	1,417	1,531
Finance, insurance	996	1,344	1,809	1,860
Service except domestic	2,137	3,174	948	1,079
Domestic	1,785	1,942	598	637
Transportation	2,655	2,273	1,536	2,007
Utilities	570	915	1,251	1,826
Manufacturing	9,078	13,335	1,377	1,764
Government	2,901	5,153	1,210	1,631

the same difference between earnings at different ages existed in 1920 as did in 1939 we find about a 3 per cent rise in earnings because of changes in age composition.[10] This 3 per cent rise is based on a 5 per cent rise because of changing age composition, 2 points of which were lost because of the increased importance of women in the labor force, women generally receiving lower incomes than men.

TABLE 5

EMPLOYMENT (1920 AND 1940) AND EARNINGS OF FULL TIME EMPLOYEES (1939) BY OCCUPATION GROUP

	Male			Female		
	Employment (000)		Median Full Time Earnings of Employees 1939	Employment (000)		Median Full Time Earnings of Employees 1939
	1920	1940		1920	1940	
Proprietors and managers exc. trade	1,314	1,664	$2,600	78	193	$1,450
Clerks	3,491	4,785	1,500	2,191	3,248	1,000
Skilled workers	5,469	5,089	1,600	102	87	1,200
Semi-skilled workers	4,371	6,190	1,250	2,260	3,072	950
Laborers exc. farm	5,819	3 648	1,000	200	100	750
Servants	675	1,300	1,000	1,570	2,346	400

---

[10] The age distributions are based on those for non-agricultural gainful workers other than those in trade as given in the 1920 Census of Population, Vol. IV, p. 376, and the 1940 Census of Population, Vol. III, Pt. I, pp. 197, 199, Earnings by age and sex are from 1940 Census of Population, *Wage and Salary Incomes in 1939*, p. 107.

It is not possible simply to add the 1.6 point rise in earnings because of changing industrial composition, the 3.0 per cent rise because of changing age composition and the gain, say of 0.5 per cent, because of varying occupational composition. This is so partly because the figures may not be exactly compared, partly because the changing industrial composition overlaps the age change, and so on. We must therefore be content with the understanding that of the 22 per cent gain in full time earnings something upwards of three-quarters arose because of a rise in earnings on a given job and a relatively small amount derived from a rising level of skill in the labor force or the rising level of maturity. (With the essential addendum that part of the rise which did take place was nullified by the increased role of women in the labor force.[11])

II

Comparable estimates of full time equivalent earnings by industry for the years 1919–1943 are given in Tables 6 and 7. In terms of full time earnings as, Chart I emphasizes, the various manufacturing industries retained their general position with reference to one another over the period 1919–23 to 1939–43. Petroleum and coal products manufacturing was the highest paying industry in both quinquennia while tobacco products was the lowest in one, next to the lowest in the other. The sharpest shift was that of printing, which fell from 2nd to 7th place, while nonferrous manufacturing rose from 7th to 4th. The similarity of structure is likewise evident in data for major industrial groups (Chart II), although it is nowhere near as neat.[12]

In general the per cent changes in income were related to the changes in employment as Chart III suggests for manufacturing industries and as the freehand curve in Chart IV less reasonably suggests for major industrial groups. It is clear that there is a close relationship between changes of income and employment in manufacturing, a much rougher

---

[11] Inasmuch as the earnings data used for the occupational and age calculation were from the 1940 Census it might be noted that by also using median earnings data by industry for 1939 from the census (*Wage and Salary Income in 1939*, pp. 146–147) in place of the 1939–43 full time equivalent averages, one would get a rise one-half of one per cent greater.

The reader may not unreasonably be surprised by the small gains attributable to occupational and industrial shifts. Several factors might be noted. (1) These data apply only to nonfarm employees; the rising importance of service industries and the declining importance of agriculture is therefore irrelevant. (2) A rising proportion of women in the labor force has tended to limit the rise in the average wage. (3) Because in point of fact a rise ascribed to a change in industrial composition only came about through the operation of changes in age and occupations as well it will not do to emphasize one factor to the exclusion of the others.

[12] A more extended discussion appears in the writer's "Wage Structures," *Review of Economic Statistics*, November, 1947.

CHART I.
EARNINGS AND EMPLOYMENT BY MANUFACTURING INDUSTRY,
1919-23 AND 1939-43

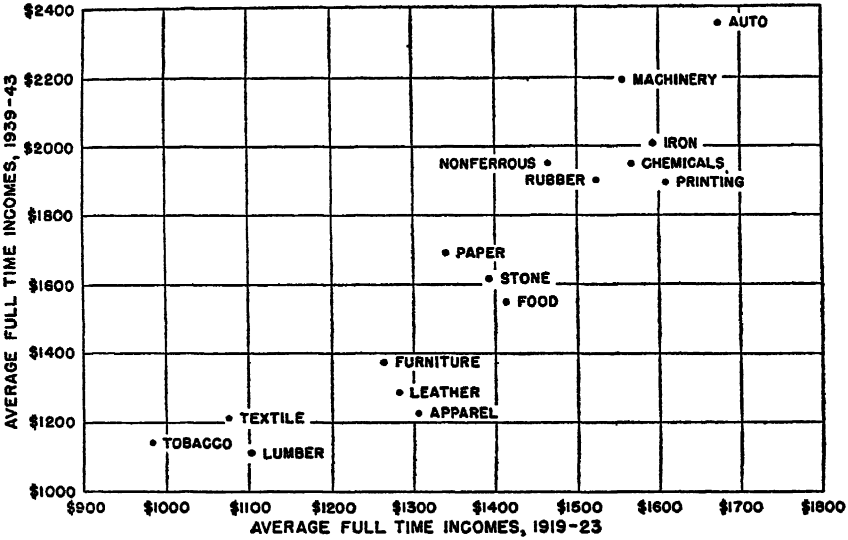


CHART 2.
EARNINGS AND EMPLOYMENT BY MAJOR INDUSTRY GROUPS,
1919-23 AND 1939-43

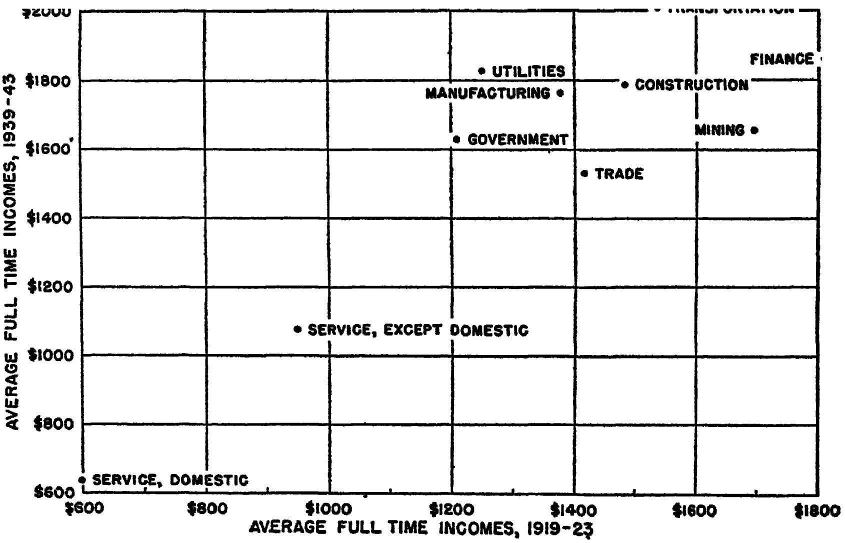


CHART 3.
PERCENT CHANGE IN FULL TIME EQUIVALENT EMPLOYMENT AND INCOME, 1919-23 AND 1939-43, BY MANUFACTURING INDUSTRY

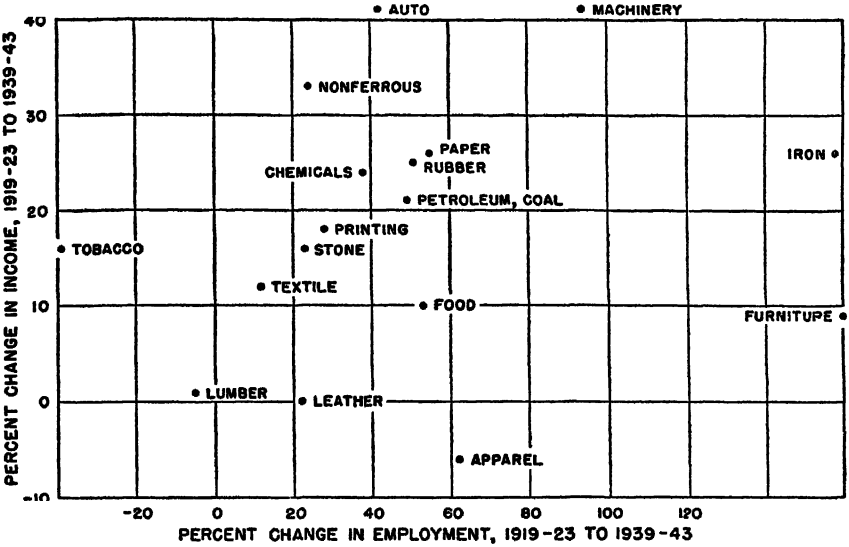
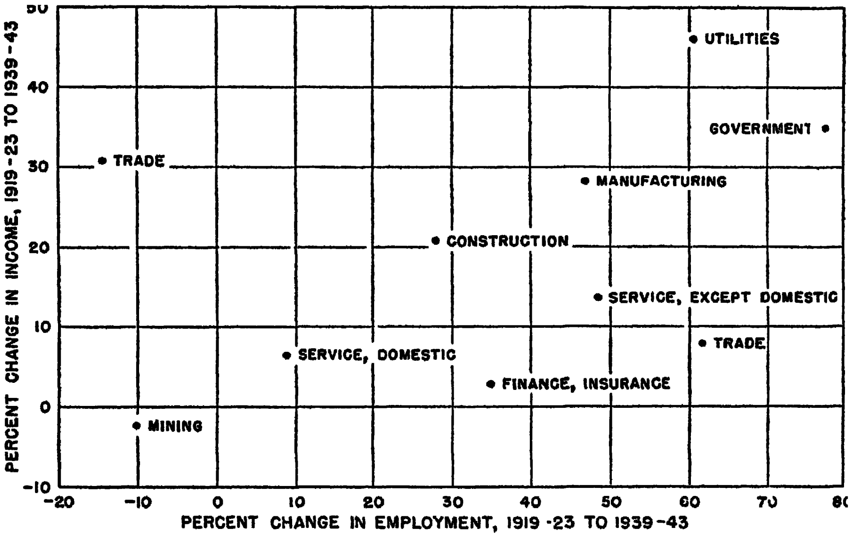


CHART 4.
PERCENT CHANGE IN FULL TIME EQUIVALENT EMPLOYMENT AND INCOME, 1919-23 AND 1939-43, BY MAJOR INDUSTRY GROUPS



Erratum: For TRADE in upper left quadrant (30 7,-14 1) read TRANSPORTATION.

one for major industries.[13] Whether this contrast arises from the fact that productivity changes in manufacturing industries tended to occur in unison and proportion unlike those in nonmanufacturing, or whether it arises from still other factors cannot be sufficiently dealt with in the compass of the present article.

TABLE 6

AVERAGE ANNUAL EARNINGS PER FULL TIME EMPLOYEE

Major Industry Groups

Year	Mining	Contract Construction	Manufacturing	Wholesale and Retail Trade	Finance, Insurance and Real Estate	Transportation	Communication and Public Utilities	Domestic Service	Service except Domestic
				(in	dollars)				
1919	1,427	1,394	1,304	1,397	1,578	1,403	1,087	510	1,103
1920	1,758	1,709	1,532	1,436	1,757	1,694	1,267	578	1,201
1921	1,807	1.386	1,345	1,361	1,865	1,571	1,300	620	1,201
1922	1,657	1.305	1,295	1,419	1.939	1,501	1,230	633	1,197
1923	1,831	1,614	1,410	1,472	1,904	1,513	1,313	647	1,219
1924	1,738	1,621	1,431	1.462	1,954	1,529	1,383	668	1,242
1925	1,620	1,656	1,454	1,539	2,010	1,553	1,390	681	1,252
1926	1,655	1,664	1,478	1,583	2.022	1,569	1,432	684	1,267
1927	1,630	1,707	1,503	1.522	2,033	1,579	1,443	690	1,324
1928	1,570	1,717	1,535	1,546	2,059	1,609	1,474	688	1,313
1929	1,526	1,674	1,543	1,597	2,090	1,642	1,474	701	1,371
1930	1,424	1,526	1,488	1,568	2,001	1.610	1,497	650	1,373
1931	1,221	1,233	1,369	1,497	1,886	1,549	1,514	561	1,326
1932	1,016	907	1,150	1,318	1,687	1.373	1,438	477	1,210
1933	990	869	1,086	1,187	1,591	1,334	1,351	442	1,119
1934	1,108	942	1,153	1,232	1,635	1 393	1,426	455	1,129
1935	1,154	1,027	1,216	1,281	1,668	1,492	1,486	467	1,150
1936	1,263	1,178	1,287	1,299	1,747	1,582	1,522	487	1,181
1937	1,366	1,278	1,376	1,356	1,819	1,644	1.601	536	1.215
1938	1,282	1,193	1,296	1,357	1,762	1,676	1,674	506	1,220
1939	1,367	1,268	1,363	1,365	1,761	1,723	1,692	520	1,235
1940	1,388	1,330	1,432	1,391	1,754	1,754	1,718	533	1,240
1941	1,579	1,638	1,653	1,401	1,805	1,888	1,766	578	1,291
1942	1,795	2,194	2,023	1,626	1,918	2,181	1,881	678	1,400
1943	2,160	2,505	2,350	1,804	2,071	2,491	2,075	876	1,567
1944	2,499	2,602	2,517	1,965	2,203	2,677	2,248	1,080	1.724
1945	2,618	2,612	2,525	2,134	2,365	2,732	2,416	1,230	1,839
1946	2,677	2,581	2,512	2,392	2,567	2,937	2,560	1,328	2,032

---

[13] The relationship between income and employment changes in manufacturing, and these changes in relation to productivity gains, has been ably discussed in Solomon Fabricant's *Employment in Manufacturing, 1899–1939*, 1942, especially pp. 100–113.

III

To establish the basic series for real earnings of nonfarm employees three main constituent series are required: one for full time equivalent earnings, a second for per cent of time lost by unemployment and a final series for changes in the cost of living. The derivation of each of the required series will be discussed in turn.

Full time equivalent earnings. Series for the 1929–1940 period are available from the Denison estimates for the Department of Commerce.[14] For 1919–28 estimates have been made by Kuznets, but these estimates are not comparable with the Denison figures.[15] The series were examined for the overlap period to determine the best method of linking. For all groups the similarity of movement was obvious, but the levels differed in most instances. After a number of necessary regroupings of industries had been made,[16] the various series were then linked on the basis of the regression relationship for the period 1929–38.[17] In a number of instances the computed Commerce-level estimates were then ratio linked to the Commerce estimates for 1929 to prevent a reversal or exaggeration of the 1928–30 movement.

This 1919–43 series was then extrapolated to 1910 by the relationship between this series and that of W. I. King, in the overlap period.[18,19] This series in turn was further extrapolated to 1890 by the 1910–26

---

[14] *Survey of Current Business, National Income Supplement*, July 1947.

[15] Simon Kuznets, *National Income and Its Composition, 1919–1938* (1941).

[16] The following adjustments in Kuznets' series were made: hosiery and knit goods added to textile mill, deducted from apparel; leather footwear deducted from apparel, added to leather; rubber footwear deducted from apparel, added to rubber; heating apparatus added to iron and steel; restaurants added to trade, deducted from service.

[17] Lumber and timber basic products; printing and publishing; iron and steel were considered to be directly comparable.

The absolute differences between the Kuznets and Denison series are for the most part small enough, the similarity of movement great enough, that it is even possible to splice the series by a simple ratio link. Except for these three groups, however, regressions were used.

[18] W. I. King, *The National Income and Purchasing Power*, 1930, pp. 56, 60, 122. Bowden has revised some of the basic data which King relied on, but the use of these substitute figures makes little difference in the final estimates here presented. Cf. Witt Bowden, *War and Post-War Wages, Prices and Hours, 1914–23 and 1939–44*. BLS Bulletin 852. Additional data of Bowden's, for earlier years, appear in the *Monthly Labor Review*, September, 1940.

[19] King's figures have been adjusted by Kuznets with a view to comparability (Kuznets, *op. cit.*, Vol. II, pp. 469–474) but these adjustments do not affect our present concern. While King's implicit average compensation figure for nonfarm employees is $1,220 in 1919 and $1,479 in 1920, the estimates here presented are $1,260 and $1,448 respectively. King's figures were therefore assumed to be directly comparable with the present estimates. The 1919–43 series could, of course, have been extrapolated all the way back to 1890 by Douglas. But instead of such a procedure, based on the 8-year overlap from 1919–26, it seemed preferable to use King's figures back to 1910, thus producing an independent series which could be related to the Douglas series for the 17-year period, 1910–26.

TABLE 7 (PART 1)

AVERAGE ANNUAL EARNINGS PER FULL-TIME EMPLOYEE

Manufacturing Industries

1919-1946

Year	Food and Kindred Products	Tobacco Manufactures	Textile-Mill Products	Apparel and Other Finished Fabric Products	Lumber and Timber Basic Products	Furniture and Finished Lumber Products	Paper and Allied Products	Printing, Publish- and Allied Industries	Chemicals and Allied Products
					(in dollars)				
1919	1,319	942	1,013	1,237	1,087	1,151	1,236	1,320	1,467
1920	1,470	1,083	1,226	1,405	1,376	1,308	1,461	1,666	1,780
1921	1,448	950	1,052	1,284	1,013	1,282	1,320	1,670	1,531
1922	1,386	950	1,001	1.275	942	1,258	1,308	1,650	1,499
1923	1,437	988	1,110	1,322	1,094	1,317	1,375	1,737	1,556
1924	1.468	981	1,099	1,318	1,107	1,345	1,402	1,812	1,581
1925	1,460	1,001	1,114	1,343	1,121	1,359	1,421	1,868	1,618
1926	1,484	1,006	1,119	1,331	1,135	1,385	1,452	1,964	1,627
1927	1,510	976	1,151	1,364	1,155	1,386	1,470	1,969	1,644
1928	1,516	965	1,141	1,362	1,164	1,379	1,501	2,011	1,686
1929	1,503	979	1,155	1,361	1,172	1,398	1,514	2,010	1,673
1930	1,489	985	1,096	1,265	1,156	1,310	1,487	2,011	1,647
1931	1,451	908	1,039	1,162	1,010	1,196	1,404	1,943	1,608
1932	1,303	787	847	941	787	962	1,028	1,740	1,419
1933	1,204	725	829	900	737	900	1,143	1,599	1,312
1934	1,221	750	883	987	791	948	1,186	1,644	1,341
1935	1,253	778	926	1,016	833	988	1,235	1,698	1,385
1936	1,290	817	952	1,013	911	1,074	1,313	1,702	1,455
1937	1,351	883	994	1,025	963	1,123	1,403	1,722	1,559
1938	1,331	870	926	999	940	1,102	1,359	1,697	1,621
1939	1,372	916	960	1,025	956	1,138	1,414	1,718	1,611
1940	1,385	1,000	986	1,022	934	1,158	1,458	1,764	1,723
1941	1,472	1,117	1,159	1,159	1,026	1,304	1,646	1,852	1,893
1942	1,650	1,240	1,385	1,330	1,204	1,514	1,850	1,973	2,139
1943	1,879	1,431	1,556	1,595	1,449	1,743	2,076	2,158	2,386
1944	2,044	1,580	1,681	1,788	1,564	1,892	2,254	2,376	2,608
1945	2,176	1,693	1,814	1,944	1,617	1,983	2,363	2,577	2,683
1946	2,392	1,798	2,037	2,168	1,781	2,203	2,547	2,871	2,752

relationship between it and the Douglas series.[20]

Time lost in unemployment. Given a series for full time earnings our next requisite is a parallel one for time lost in unemployment by nonfarm employees so that the earnings series can be transformed from a full time equivalent one into an actual average series. Total unemployment estimates for the 1929–40 period have been made available recently by the Bureau of Labor Statistics.[21] Comparable but unpub-

---

[20] Douglas, *Real Wages*, Table 147, p. 392, average earnings in all industries excluding farm labor. Douglas notes that his series represents the movement of earnings of 73 per cent of those who worked for wages or salaries in 1920. (*Op. cit.*, p. 389.) Without attempting to revise his estimates on the basis of later data it is nonetheless clear that his series includes the bulk of nonfarm workers and hence can be readily used for extrapolating a nonfarm earnings series.

[21] Cf. The writer's (Bureau of Labor Statistics, Occupational Outlook Division, July 1945), *Preliminary Estimates of Labor Force, Employment and Unemployment, 1929–1940.*

TABLE 7 (PART 2)

Products of Petroleum and Coal	Rubber Products	Leather and Leather Products	Stone, Clay, and Glass Products	Iron and Steel and Their Products (including Ordnance)	Non-ferrous Metals and Their Products	Machinery and Transportation Equipment (except Automobiles)	Automobiles and Automobile Equipment	Total All Manufacturing	Year
				(in dollars)					
1,671	1,446	1,231	1,271	1,599	1,360	1,493	1,531	1,304	1919
2,224	1,759	1,377	1,528	1,882	1,596	1,700	1,818	1,532	1920
1,808	1,410	1,273	1,397	1,489	1,414	1,543	1,668	1,345	1921
1,725	1,442	1,258	1,308	1,377	1,434	1,446	1,617	1,295	1922
1,794	1,530	1,274	1,460	1,622	1,516	1,605	1,735	1,410	1923
1,782	1,561	1,264	1,536	1,633	1,580	1,629	1,720	1,431	1924
1,804	1,542	1,278	1,519	1,652	1,588	1,658	1,795	1,454	1925
1,821	1,573	1,287	1,519	1,667	1,639	1,678	1,786	1,478	1926
1,832	1,619	1,303	1,546	1,677	1,636	1,717	1,809	1,503	1927
1,858	1,633	1,296	1,568	1,738	1,675	1,754	1,864	1,535	1928
1,844	1,597	1,327	1,557	1,740	1,665	1,756	1,813	1,543	1929
1,904	1,563	1,215	1,525	1,640	1,554	1,715	1,571	1,488	1930
1,810	1,392	1,152	1,386	1,410	1,455	1,507	1,455	1,369	1931
1,619	1,191	970	1,167	1,044	1,177	1,283	1,234	1,150	1932
1,505	1,137	950	1,071	1,073	1,132	1,245	1,170	1,086	1933
1,513	1,248	1,017	1,088	1,166	1,209	1,320	1,314	1,153	1934
1,587	1,358	1,043	1,171	1,295	1,277	1,400	1,489	1,216	1935
1,629	1,472	1,045	1,262	1,446	1,361	1,520	1,600	1,287	1936
1,833	1,526	1,085	1,357	1,591	1,492	1,658	1,672	1,376	1937
1,863	1,457	1,017	1,303	1,359	1,402	1,538	1,653	1,296	1938
1,852	1,548	1,038	1,359	1,549	1,521	1,653	1,762	1,363	1939
1,954	1,583	1,041	1,393	1,643	1,594	1,767	1,934	1,432	1940
2,113	1,778	1,236	1,554	1,923	1,824	2,091	2,243	1,653	1941
2,410	2,116	1,447	1,771	2,284	2,235	2,592	2,880	2,023	1942
2,806	2,478	1,659	2,024	2,637	2,581	2,862	2,978	2,350	1943
3,046	2,699	1,831	2,174	2,781	2,724	3,022	3,103	2,517	1944
3,092	2,715	1,969	2,252	2,806	2,741	2,994	2,984	2,525	1945
3,180	2,826	2,123	2,394	2,687	2,710	2,808	2,796	2,512	1946

lished figures for 1920–29 are likewise available.[22] These data constitute a logical starting point.

Before this series can be used as a measure of time lost by nonfarm employees, however, we must consider to what extent the unemployment of farm hands and self-employed persons is included in that series. If it can be established that unemployment in those groups can be balanced off against the failure of the series to allow for the full measure of time lost by nonfarm employees, then the series can be used as is.

[22] Described by the writer in a forthcoming BLS technical memorandum.

On balance it may be assumed that the nonfarm self-employed contributed nothing to the unemployment totals. The small number of nonfarm self-employed in April 1940, for example—roughly 0.2 out of 4.4 million experienced workers seeking work—was more than compensated for by the number of disemployed workers who, turning to self-employment, had disguised their unemployment in that form.[23]

Agricultural workers—operators and hired hands—constituted about 4 per cent of total unemployment in April 1930 and about 8 per cent of the experienced unemployed in March 1940.[24] More than compensating for the inclusion of agricultural unemployment is the exclusion of two groups of unemployed. First is the group of women and older workers who, after being without work for some time, withdraw from the labor force.[25] Second is the group of workers who migrated to rural areas because of lack of work in nonfarm pursuits, the relationship between farm-nonfarm migration and the level of employment being well known. It may be added that even if we were to ignore these two groups and allow in full for agricultural unemployment, the resultant estimate for 1940 would be 18.3 per cent of time lost rather than 19.9.

Granted these various considerations, therefore, it was decided to take the level of unemployment shown by the BLS estimates as equivalent to the level of unemployment among nonfarm employees.[26] The per cent of time lost is computed as the ratio of unemployment on the one hand to employed nonfarm employees plus unemployed on the other.

Given unemployment estimates for 1920–43 we must next derive serviceable estimates for prior years. National unemployment enumerations were conducted and data published in connection with both the

---

[23] 1940 Census of Population, *Employment and Personal Characteristics*, Table 11, with a deduction for unemployed farm operators (1940 Census of Population, *Occupational Characteristics*, Table 6). A discussion of the relatively small decline in self-employment appears in the BLS, *Estimates of the Non-agricultural Self-Employed, 1929–1940* (1945).

[24] 1930 Census of Population, *Unemployment*, Vol. I, pp. 53–54, Class A and B unemployment. 1940 Census of Population, *The Labor Force*, Vol. III, Table 59; data in Table 1 for rural farm areas allow us to take account of new workers and public emergency workers but necessarily include some nonfarm workers. The rural farm ratio is 11 per cent. It should be noted that some 20 per cent of the experienced persons seeking work in March 1940 who reported their last occupation as farmer or farm operator did not report their usual occupation in the same group; this contrasts with the over-all male average of 10 per cent. (1940 Census of Population, *Usual Occupations*, Tables 11 and II.) There is likely to be some inflation in agricultural unemployment, therefore, precisely because workers have exported their unemployment from cities to rural areas, attempting to become farm workers in the process. On the other hand, some unemployed farm workers undoubtedly sought work in the cities.

[25] Gladys Palmer has called attention to "the substantial increase we found in Philadelphia in the number of households reporting no employable member as the depression continued." (With respect to the 1929–40 period, the point is discussed at greater length in a forthcoming BLS Bulletin.)

[26] Given a continuing substratum of disguised unemployment these percentages are probably somewhat too small. More important for present purposes is the fact that the undoubted gain in the importance of such unemployment during the depression is not reflected in the estimates of time lost. As a result the gain in nonfarm earnings over the years is exaggerated.

1890 and the 1900 Census of Population. However it does not now appear possible to make any use of the 1890 estimates.[27] An estimate for 1900, however, can be computed from the Census data after certain adjustments are made.[28]

(1) The first such adjustment is the exclusion from both the gainful worker and the unemployment totals of those workers who do not come within the purview of the post-1940 "labor force" concept. For male workers it is possible to make such exclusions with at least approximate accuracy, utilizing for the purpose the 1901 Cost of Living survey.[29]

(2) Since the 1900 Census distributes the unemployed into three unemployment duration groups it is then necessary to estimate the average duration for each group before total time lost in unemployment by males can be estimated. As the 1940 Census data indicate it is not possible to secure an accurate average for a group as large as 7 to 12 months of unemployment merely by taking the mid-point. However, the 1901 Cost of Living survey gives a distribution into 23 duration groups and it is thus possible to calculate an average duration of unemployment for men in the nonagricultural labor force in 1899–1900.[30]

(3) In step one the appropriate percentage of total male unemployed nonagricultural gainful workers in 1899–1900 was deducted in order to ensure comparability with later data, using proportions from the 1901 Cost of Living survey. In step two medians were computed for each of the three unemployment groups after they had been thus adjusted. In step three a simple multiplication gave the estimated total weeks lost in 1899–1900 by the male nonagricultural labor force. After corresponding adjustments in the gainful worker figures to give total labor force weeks for this group it was possible to compute the per cent of time lost by it—4.93.[31]

---

[27] The 1890 data, unlike those for 1900 and later years, apply chiefly to unemployment at the worker's primary job rather than to total unemployment. Furthermore there is clear evidence that the 1900 Census enumerators operated with a better understanding of what they were seeking, and hence obtained estimates which were generally higher than those of 1890 merely because the enumeration technique differed. 1890 Census of Population, Part II, p. cxxvi. 1900 Census of Population, *Occupations*, pp. ccxxvi–ccxxxiii.

[28] 1900 Census, *op. cit.*, pp. 76–77.

[29] Eighteenth Annual *Report* of the Commissioner of Labor, 1903 (1904), pp. 290–291. It was assumed that those workers whose unemployment was ascribed to the following causes would have been excluded by 1940 definitions: old age, sickness, strike or vacation. On the other hand those for whom any of these causes was reported in combination with another cause were included. This was done because enumerative procedures in current use undoubtedly include some such workers in practice, and definitely include many who are temporarily ill and without a job.

[30] Durations, of course, were computed for the groups of workers remaining after the excluded workers had been deducted. Data on those unemployed for 52 weeks were taken from page 26 of the *Report*, while all other data appear on pages 290–291.

[31] Both teachers and manufacturers were excluded from these computations since according to current practice few if any would be reported as unemployed.

(4) This percentage is very substantially below an estimate of 9.4 per cent to be derived from the Cost of Living *Report*.[32] A major reason for the divergence is the fact that according to the Report 50 per cent of male urban heads of families had some unemployment during a year ending in 1901 while the Census figure for males except in agriculture was 23 per cent.[33]

It seems reasonable to believe that in this respect the Cost of Living report is the more accurate measure. In such surveys, where income was reconciled with expense on a detailed basis of reporting, there was less likelihood that short period unemployment would be forgotten than there was in the general purpose population enumeration, which could necessarily have little check on the solidity of the unemployment reports in the absence of questions on income or related factors. It is therefore assumed that the Census figure is an underestimate. An upward adjustment was made by doubling the number in the lowest duration group, exclusive of those in agriculture, teachers and manufacturers. This step raised the 1899–1900 percentage to 8.3—or somewhat below the adjusted 1901 figure of 9.4.

This divergence was ignored because the unemployment figures reported to the Census for nonfarm women are likely to be too high.[34] Combining the estimate of 8.3 per cent with the parallel figure for female workers an over-all percentage of 7.5 was finally secured.[35] This percentage then serves as a benchmark figure for 1900, comparable with the BLS 1920–40 estimates.

If it were possible to utilize the well known estimates of Douglas for the intervening years then a complete unemployment series could be calculated. Douglas's absolute estimates, however, cannot be used as is. If one follows his procedure of comparing the labor supply (estimated from Census of Population data) with employment (estimated from payroll sources) but uses the revised labor supply data of Daniel Carson and the revised employment estimates of the BLS one arrives at sharply different results for a test year.[36] While the absolute figures in Table 8 may not be compared the percentages can be.

---

[32] The percentage to be derived directly from the Report is 6.8. However, this must be adjusted to allow for self-employed and for single persons—both groups being included in the Census data. In 1940 there were 7 male nonagricultural employees to 1 employer or own-account worker. (1940 *Population Census*, III, p. 1.) Though trade, service and construction together made up a much smaller fraction of the nonfarm total in 1940—as Carson's data demonstrate—we can nonetheless use the 1940 proportion and thus intentionally underestimate the spread between the Report and the Census estimates. Assuming a 0 rate of unemployment for the self-employed we find the adjusted 1901 rate is 6.0. However by assuming that the unemployment rate for urban single workers in 1900 was the same proportion of the rate for urban married workers in 1900 as it was in 1940 (1940 Census of Population, III, p. 22) the head of family rate as adjusted from the 1901 Report was used to construct a single workers rate. The two were then weighted by the distribution of nonagricultural males by marital status (1900 Census, *Occupations*, p. 52) with the result that the 6.0 figure rises to 9.4 per cent.

[33] Report, p. 287. *Occupations*, pp. 7, 76.

[34] A comparison of unemployment rates by occupations shows that female rates for various occupations were above male rates in 1900, instead of almost uniformly below them as in 1930 and 1940. Since women are not as continuously in the labor force as men are, what is reported as "unemployment" in some instances is really "non-employment," or time not in the labor force. (This factor was adjusted for in the male data on the basis of the 1901 Report.)

[35] Data for agricultural unemployment are excluded since the chief reasons which dictated their inclusion in later years do not apply to the 1900 data. Their inclusion would only reduce the 1900 earnings estimate by $21.

TABLE 8

UNEMPLOYMENT IN SELECTED INDUSTRIES IN 1920 AS ESTIMATED FROM DOUGLAS'S, GIVENS' AND REVISED DATA

Manufacturing and transportation	Labor Supply	Employment	Per cent Unemployed
Carson—BLS	14,401	14,492	−0.6
Douglas—Unadj.*	10,732	10,628	1.0
Givens†	15,418	14,761	4.3
Manufacturing, transportation, construction			
Present estimate	15,521	15,348	1.1
Douglas‡	12,221	11,478	6.1

* Manufacturing plus steam railroads plus street railways. Data prior to arbitrary adjustment to Givens' unemployment percentage

† Includes hand trades in labor supply and possibly in employment as well.

‡ After adjustment of manufacturing and transportation data to Givens' unemployment percentage.

The percentage of manufacturing and transportation combined is negative and by any reasonable interpretation is valueless. When construction is added the percentage becomes a positive one, but its absolute value may still be in question. However the series may be reasonably indicative of the changing movement of unemployment. Given benchmark data for 1900 and 1920 the Douglas data may be used for interpolation. A regression of the present series for per cent of time lost 1920–26 against that of Douglas' series (for manufacturing, mining, construction and rail transport) was therefore computed.[37] This indicated a sufficiently close relationship so that the present series could be carried back to 1890. The resultant estimated series was then used to interpolate between the 1900 benchmark and the 1920 estimate.[38] The consideration of such related materials as the trend of deflated gross product, unemployment among union labor in New York and the qualitative data in Thorp's *Business Annals* tended to confirm the verisimilitude of the movement thus indicated.[39]

---

[36] Douglas' data from *Real Wages*, Chapter 24. The Carson estimates are from his *Industrial Composition of Manpower in the United States, 1870–1940* (November 1946).

The labor supply in construction was not estimated from the Carson data since he computes the number of construction laborers in 1920 by applying the 1930 ratio of laborers to number of workers in selected construction trades. However an unduly high proportion of construction workers in 1920 reported themselves in the skilled trades, presumably calling themselves carpenters, painters, etc., as a result of work in war construction. The labor supply was therefore estimated by applying to the BLS employment estimate the unemployment percentage used by Douglas (p. 451).

The Givens' estimates are from *Recent Economic Changes* (1929), Vol. II, pp. 466–478. For most of his 1920 unemployment percentages Douglas uses the Givens' estimates.

NICB estimates are likewise available for this period, and indeed back to 1900. *Economic Record*, March 20, 1940. These data for the 1900–1920 period, however, seem of questionable value. The 1920 unemployment percentage, for example, is the resultant of a combination of percentages for each industry—the percentages secured by interpolating between a 1900 and a 1930 percentage, according to a communication from the NICB. The 1920 figure, however, is likely to have been less than either, and this distortion serves to distort the entire 1900–30 level of the estimates.

The per cent of time lost for 1940–46 can be readily computed from the Bureau of Census data, thus providing a full series for the period 1890–1946.[40] By applying this series to that previously computed for full time earnings we secure the final estimates for the actual earnings of nonfarm employees. The nominal earnings series together with one adjusted for changes in the cost of living appears in Table 2.[41,42]

It may not be without interest to compare the movement shown in Table 2 with Douglas' figures for the period when the two full time earnings series are independent: 1910 to 1926. Thus Douglas' full time earnings figure for nonfarm wage earners rises 126 per cent from $652 in 1910–14 to $1,473 in 1926.[43] The present estimate, is necessarily higher in 1910–14, being for wage earners plus salaried workers, but its rise to $1,458 in 1926 shows an increase of 108 per cent. In part the difference in movement comes about because Douglas' series, excluding as it does construction, trade, finance and other groups, is not adapted to allow for the effect of changes in the industrial, occupational and sex composition of the labor force which result when these groups are included. But discussion of the difference should not be allowed to obscure the fact that the present estimate of change over this period is within 15 per cent of Douglas'—a variance easily explained by the factors noted.

---

[37] Douglas, *op. cit.*, p. 460, presents estimates for the 4 specified industries for 1897–1926. His series was extrapolated to 1890 by its relationship, over the 1897–1926 period, to his series for unemployment in manufacturing and transportation (*ibid*, p. 445).

[38] The 1900 value estimated from the regression was 7.9 per cent, as compared to 7.5 from the Census materials.

[39] The relationship between unemployment percentages and per cent deviations of gross national product from trend is dissussed in the forthcoming BLS memorandum on the 1920–29 estimates. As indicated there the relationship is not such that one can assume the relationship between the two series during the 1920's had the same slope as that in earlier years. Hence it could not be assuredly used for extrapolation. However the fact that the slope was like the 1920's, and not the 1930's, did tend to confirm a relationship otherwise arrived at.

For the New York data, cf. BLS Bulletin No. 109, *Statistics of Unemployment and the Work of Employment Offices*, 1913, p. 18.

[40] Bureau of the Census, *Labor Force, Employment and Unemployment in the United States, 1940–1946* (1947).

[41] The cost of living index is the regular BLS series extrapolated by a regression on Douglas' series (*op. cit.*, p. 60). It would have been possible to make some allowance for the BAE series on prices paid by farmers for food, prices paid for house furnishings etc.—giving these some weight as representative of prices paid by rural nonfarm workers. However, the aggregate weight of this group is relatively small and the differences in movement between BAE and BLS price series relatively restricted. Therefore the BLS series was used as is.

[42] The actual earnings series developed here was inflated by the ratio of total compensation to total wages and salaries for the 1920–43 period. The relevant 1920–29 data appear in Kuznets, *National Income and Its Composition*, Vol. 1, p. 216. The later figures are from the *Survey of Current Business*, July 1947. For the earlier years no adjustment is necessary, King's figures for total compensation having been used initially.

The change in unemployment percentages over the 1910–1926 period is, of course, almost identical in magnitude, the movement of the present series being based on Douglas' for most of these years. (Douglas'—for manufacturing, mining, transport and construction—falls by 22 per cent, while the present one for all nonfarm declines by 21 per cent.) It is hardly surprising, therefore, that the net change in real earnings which Douglas shows for those 4 industries is very similar to that shown for all employees in the present estimates, 133 per cent in one instance, 112 per cent in the other.[44,45]

---

[43] Douglas, *op cit*, Table 147, p. 390.

[44] *Ibid.*, pp. 460, 468. With further reference to the similarity, in a general sense, it should be noted that materials on full time earnings and employment drawn upon for use in this present study derive in no small part from the prodigious and invaluable labors of Douglas.

[45] The absolute levels of unemployment percentages are, of course, not comparable: Douglas' apply to only part of all wage earners, albeit the major part. As he has emphasized there is no separate labor supply for manufacturing since there is an interchange between its workers and those in construction, transport, etc. (*ibid*, p. 602). It follows, therefore, that unemployment percentages should be computed over as wide a base as possible—preferably all nonfarm workers, as is done in the present study.

DIRECT DETERMINATION OF COMPASS SETTINGS FOR PROPORTIONATE AREA PIE-CHARTS*

HERMAN LASKEN
Inglewood, Calif.

This note presents a chart which can be used by draftsmen to determine directly from raw area data compass settings for plotting proportionate area circles without any of the usual square-root and proportioning computations.

AVOIDING here any discussion of the validity, from a purely statistical standpoint, of proportionate area pie-charts as a means of presenting statistical data, the fact remains that such charts are frequently used and that their preparation involves some computations. Regardless of the order in which the computations are made, the determination of the compass settings which will achieve proportionate areas of circles requires extraction of square roots and division of all such square roots by the square root of the circle taken as the standard of measurement for the particular chart.

The accompanying chart is designed to eliminate all such computaions in the drafting of pie-charts. Following the order of the instructions which are included on the chart:

Step 1 reduces the given area data to two digits as a maximum, which is more than sufficient accuracy for this type of chart.

Step 2 amounts simply to a square-root computation. The heavy curved line is the curve of square roots corresponding to the area scale, with the vertical scale set at values of zero to ten. (The scale is omitted from the chart as it is unnecessary for the purpose of the chart and simply adds confusion.)

Steps 3 and 4 determine the compass setting for the largest circle on the chart and automatically select the oblique lettered line to be used for all remaining compass settings.

Step 5 instructs the draftsman to draw his biggest circle, with his compass as set in step 4.

Step 6 provides for similar settings and plottings for the remaining areas. All settings are to be made from the zero line to the oblique line determined as above.

The original chart is a ten-inch square, principally for convenience of use. However, its use is not affected by changes in size.

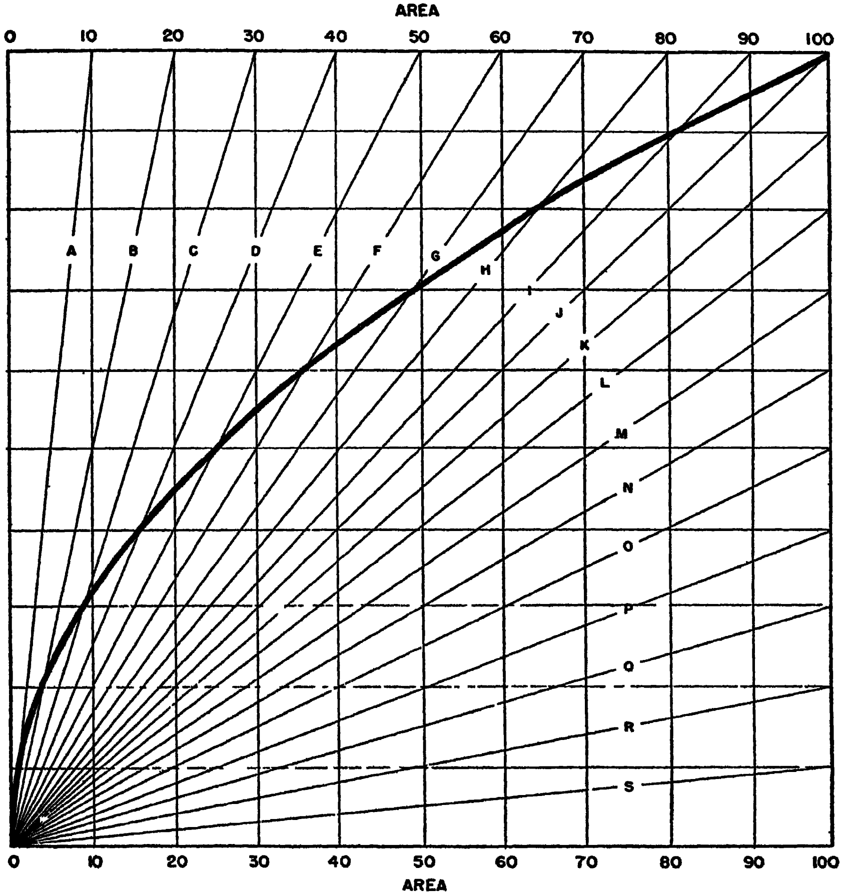
---

* The term "pie chart" as used in this note, refers to circle charts, both whole and segmented.

STANDARD CHART FOR DIRECT DETERMINATION OF COMPASS SETTINGS FOR PROPORTIONATE AREA PIE-CHARTS

INSTRUCTIONS FOR USE

1. POINT OFF GIVEN AREA DATA TO TWO PLACES FOR THE LARGEST VALUE.
2. DETERMINE POINT ON CURVE CORRESPONDING TO LARGEST AREA.
3. SET ONE POINT OF COMPASS ON ZERO LINE HORIZONTAL TO POINT SO DETERMINED.
4. SET SECOND COMPASS POINT ALONG SAME HORIZONTAL ON LETTERED LINE NEAREST RADIUS OF LARGEST CIRCLE DESIRED IN VIEW OF GENERAL SIZE AND LAYOUT OF CHART.
5. NOTING LETTERED LINE SO DETERMINED, PLOT CIRCLE ON CHART WITH COMPASS AS SET.
6. ALL OTHER SETTINGS ARE THEN MADE FROM THE ZERO LINE TO THE SAME LETTERED LINE ALONG HORIZONTALS DETERMINED FROM THE POINTS ON THE CURVE CORRESPONDING TO THE GIVEN AREAS.



Experience with the chart shows that most draftsmen require a single demonstration of its use. Once familiar with it, they use it with facility and eliminate both the usual computations and the possibility of error inherent in such computations.

PROFILE GRAPHS

JOHN V. SPIELMANS
Marquette University

This article is concerned with certain graphical devices, called "Profile Graphs," for the integrated presentation of data of several dimensions. The several types of such profiles (two-and three-dimensional, simple and composite, statistical and historical) are explained; their advantages and disadvantages discussed; and illustrations given for the several types.

THIS ARTICLE is concerned with some graphical devices for the presentation of statistical data of several dimensions. Although the device is not new[1] we feel justified in discussing it here, for two reasons. On the one hand, the method, its considerable merit notwithstanding, seems little known or at least, little used.[2] On the other hand our analysis of the types of relationships which lend themselves to this kind of presentation serves, we believe, to widen the scope of the method.

We begin with a brief discussion of the *nature of multi-dimensional relations*. By "dimensions" we mean data of distinct denomination which combine by multiplication into a further significant datum. A two-dimensional relation is thus characterized by three data, one of which is the product of the other two: $A = n \cdot m$; (or inversely, one of which is the quotient of the other two: $m = A/n$); a three-dimensional one by four data one of which is the product of the other three: $A = p \cdot q \cdot m$, and so on.

Many basic relations in the field of economics, as well as in other fields, fall into this class. To enumerate a few: value equals quantity times price; payrolls equal number of employees times average pay; daily wages equal hourly rates times hours per day; man-days of work (or of strike) equal number of workers times average days of work (or of strike); passenger-miles of traffic equal number of passengers times average mileage, and so on.

This variety of multi-dimensional relations can generally be reduced to a common type. In all such instances the data refer to classes of elements pertaining to given time intervals. In two-dimensional relations the one factor, say n, represents the number of elements in the class during the time interval in question; (number of unit goods produced, or sold, or consumed; number of acres planted; number of persons doing a certain thing; number of certain events occurring per year, or per month); while the other factor m gives a numerical attribute of the elements (prices of goods, productivity of workers, or of acres; duration of actions or of events). The product $A = n \cdot m$ then represents the total value of the attribute pertaining to the class of n elements as a whole (total value; total product; total earnings; total man-days, etc.)

---

[1] K. Karsten in *Charts and Graphs*, New York, 1925, p. 613 ff. describes certain aspects of it under the name Area Bar Charts, although, as we believe, without full appreciation of their significance.

[2] The present author hit upon a scheme of this kind in trying to give an integrated graphical presentation of strike data. See "Strike Profiles," *Journal of Political Economy*, Dec., 1944, pp. 319 ff. He has not seen it used elsewhere.

In most instances, it is true, the attribute m is not constant, but varies among the n elements in the class. In that case the aggregate value of A is obtained, not simply by multiplying, but by the summation of the m-values.

$$A = \sum m_i$$

In terms of this aggregate, however, we can define an "average" or "mean value" of m:

$$\bar{m} = \frac{1}{n} \sum_{i=1}^{n} m_i = \frac{A}{n},$$

wherefore A can again be presented as a product of two factors, $A = n \cdot \bar{m}$.

Three-dimensional relations can similarly be reduced to a common type in terms of a class or classes of elements: the whole class in question contains p subclasses of q (or $\bar{q}$) elements each, so that the total number of elements in the class is $n = p \cdot q$. If m (or $\bar{m}$) is the numerical attribute of the elements, then $A = n \cdot m = p \cdot q \cdot m$ is the aggregate m-value of the class as a whole. (For example, p acres of average productivity $\bar{q}$ bu/acre produce $n = p \cdot \bar{q}$ bu; with the price $\bar{m}$ cts/bu, the total value of the crop for the given year is $A = p \cdot \bar{q} \cdot \bar{m}$ cts.)

GRAPHICAL REPRESENTATION OF MULTI-DIMENSIONAL RELATIONS

The customary graphical representation of more-dimensional data is by way of multiple graphs: triplets of line or bar graphs in the same or in adjacent diagrams for the three data n, m, and A. (For example for quantity, price, and value of commodities, or for employment, average earnings, and payrolls, and the like.)[3]

---

[3] Since the viewpoint in such presentations does not necessarily coincide with the one taken here such multiple graphs do not always show all, or not only, the $(n+1)$ data belonging to an n-dimensional

Such multiple graphs, however, while showing clearly the data and their time variations separately, fail to give visual expression to the dimensional relation between them.

In contrast the profile graphs represent all the data which enter into the relationship in a single, dimensionally integrated figure.

A. *Profiles of two-dimensional relations.* To represent two-dimensional data one plots the two factors n and m respectively[4] as the width and height of a rectangle, whose area then represents the product datum A. Constructing such rectangles for successive time intervals, and placing them side by side, each with its proper "time label," one obtains what we call a "statistical profile." Such profile shows in a single figure the three time series $n(t)$, $m(t)$, and $A(t)$ "dimensionally integrated," the datum $A(t)$ appearing visually as the product resulting from the multiplication of the two distinct data $n(t)$ and $m(t)$.

The economy of this presentation is due to the fact that both axes of the diagram are utilized for the two dimensions n and m of the relation in question, while none is used as a "time axis." Chronological time is merely represented by the time labels of the several rectangles.

Against the advantages of this compact and integrated presentation stand certain disadvantages. First, the numerical values of the areas are less easily read and compared than in line or bar graphs showing the values of A separately. This disadvantage can be offset largely by writing the values of A inside the rectangles. (See Fig. 1.) Secondly, the time variations, not only of the areas, but also of the widths of the rectangles stand out less clearly than on graphs showing each variable separately. Again, the profiles are at a disadvantage in comparison to

---

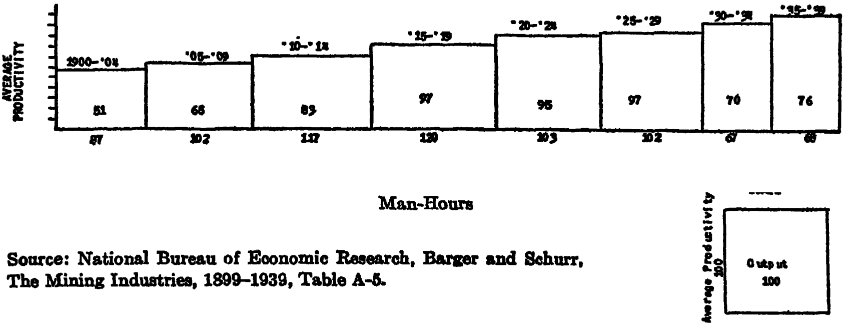
relationship. Thus the Bureau of Labor Statistics shows customarily in one diagram the indices of employment and payrolls, without the implied average earnings, while the National Industrial Conference Board (*Studies of American Wages*) shows in one diagram the indices of employment and average earnings, without the total payrolls Likewise Babson's *Business Reports* shows in multiple line graphs the quantities and prices of commodities, but not the implied total values; whereas other presentations will show doublets of line or bar graphs for quantity and total values (e.g. of crops), without the implied price. Again one finds multiple graphs of data belonging not to one, but to several two-dimensional relations. Thus the N.I.C.B. in its Weekly Chart Service, *Road Maps to Industry*, shows in one single diagram 15 line graphs of manufacturing indices, all of which can be grouped into triplets of two-dimensional data. On the other hand these and similar publications also show the exact triplets of two-dimensional data. For instance, the *Road Maps to Industry* show triplets of line graphs for indices of volume, price, and value of goods sold; of payrolls, employment, and average earnings; of number of strikes, workers involved, and average workers per strike, etc. Similarly in the *Survey of Current Business* (U. S. Department of Commerce) triplets of line graphs appear in adjacent diagrams for weekly hours, hourly earnings, and weekly earnings; for quantity, price and value of production. Since the latter part of 1937 the N.I.C.B., *loc. cit.*, separates the 15 indices of Manufacturing Changes into two groups called respectively Basic and Derived Indices. In our designation the former are the data n and A, the latter the data $\bar{m} = A/n$. While the making of this distinction draws attention to the dimensional relation between the data no visual expression is given to them in the graphs.

[4] In the following we are using m as synonymous with $\bar{m}$, unless the distinction is essential.

FIGURE 1

BITUMINOUS COAL PRODUCTION IN THE UNITED STATES, 1900–39
FIVE-YEAR AVERAGES; INDEXES, 1929 = 100

$$\text{Output} = \frac{\text{Man-Hours} \times \text{Average Productivity}}{100}$$



Source: National Bureau of Economic Research, Barger and Schurr, The Mining Industries, 1899–1939, Table A-5.

Note: Since the data presented in this figure vary on the whole but slowly a year-to-year profile would be impractical. But by presenting five-year averages a distinct and telling picture of the development as a whole results; the steadily increasing productivity per man-hour, combined with an initially strong increase, and later decrease of man-hours, resulting, after a prolonged increase, in an eventual sharp decrease in total output.

line graphs whenever the time variations of the data are only small percentages of the whole.[5] However, even under such circumstances it is often possible to obtain worthwhile profiles by constructing the several rectangles for larger time intervals. The objection has also been made that profile graphs are difficult to understand. This objection, we believe, can best be met by proper legends on the graphs showing clearly the meaning, units of measurement, and scales for all the data involved.

To summarize, the profiles are useful insofar as the advantages of the dimensional integration outweigh the several disadvantages. Moreover, one can always show in separate diagrams any one of the variables which does not stand out with desired clarity in the profile.

B. *Profiles of three-dimensional relations.* For relations of the type $A = p \cdot q \cdot m$ one can utilize the third space dimension and plot p, q, and m respectively as the length, width, and height of rectangular solids, whose volumes then represent the product A. In addition the

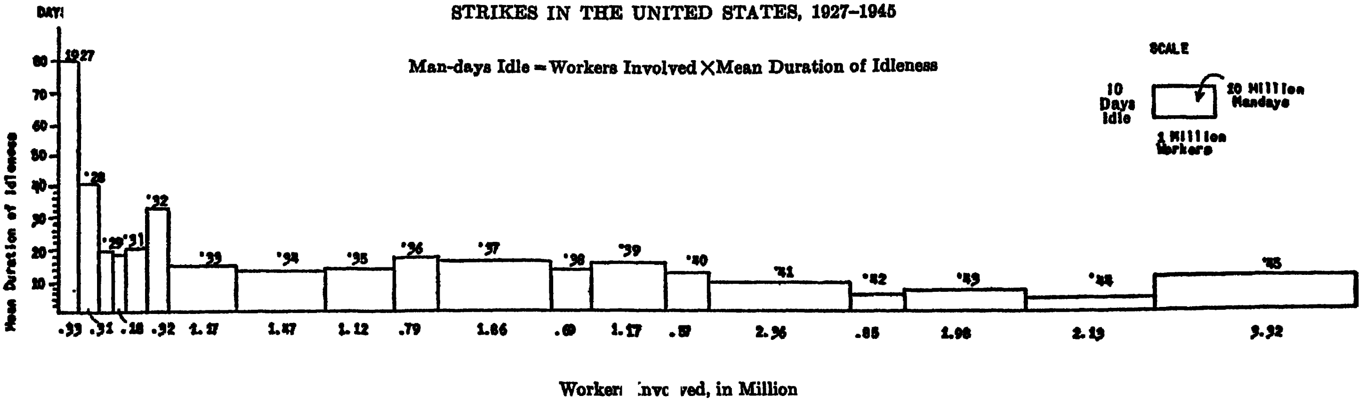
---

[5] In a line graph one can show the smallest percentage changes as large as one pleases by simply using a large enough scale, and placing the zero level into the unseen depths below the diagram. In a profile, however, where the whole amounts from zero up are represented, percentage variations cannot be made to look large when in fact they are small. Hence in such situations, where a line graph may be seen to rush dramatically up and down, a profile would offer the drab sight of a sequence of practically equal rectangles, giving, it is true, the correct information as to practically unchanging values, but unable to exhibit to the eye the small variations which do take place.

FIGURE 2

STRIKES IN THE UNITED STATES, 1927–1945

Man-days Idle = Workers Involved × Mean Duration of Idleness



Source: U

Note: In this strike profile, in which the width of the rectangles represents e number of workers involved, and the area the man-days of idleness, the height shows the average days of strike idleness per worker (mean duration of idleness). This datum although highly characteristic for the strike situation, is usually not given in strike statistics, wherefore this profile seems peculiarly revealing.* (For a discussion of the significance of this datum see Strike Profiles, *loc. cit.* pp. 320, 324, 336.) The number of strikes per year, which is not represented in this profile, can be additionally shown by "ruling." (See below, Fig. 7.)

---

* The Bureau of Labor Statistics in its strike analysis does not give the average duration of strike idleness per worker involved, but uses the "average duration of strikes." Since, at least in the U. S., the great majority of all strikes involve only rather few workers, while the very large strikes are comparatively few in number, the average duration of strikes measures essentially only the average of the small, and hence economically least important strikes. If, in forming the average duration of strikes, the individual strikes were weighted by the number of workers involved—which would seem only reasonable considering the tremendous range in magnitude, varying from a handful of men to hundreds of thousands—our "mean duration of idleness" would result. The National Industrial Conference Board in its strike analysis, based on the figures of the B.L.S., does list the "average man-days lost pe worker" (*Economic Al* , 1942–1943, p. 111). and occasionally also represents this datum graphically. (E.g. in *Road Maps to Industry*, No. 130, June 24, 1938.)

areas of the faces of the solid show the partial products $p \cdot q$, $q \cdot m$, and $p \cdot m$. Of these the first, and frequently also the second are themselves significant data, $p \cdot q = n$ as the total number of elements in the whole class, and $q \cdot m$ as the aggregate m-value per sub-class.

This presentation is very compact, being capable of showing up to six significant data dimensionally integrated in a single figure. (See below Fig. 3.) On the other hand the disadvantages noted in the two-dimensional profiles pertain *a fortiori* to the three-dimensional ones, since the volumes of solids are still less readily compared than the areas of rectangle. Hence a separate presentation of the product A may often be desirable.

To avoid these somewhat cumbersome solids one can represent three-dimensional relations in merely two space dimensions, by means of "ruled rectangles." We first construct a rectangle whose width is the partial product $n = p \cdot q$. With m for its height its area represents the three-dimensional product $A = p \cdot q \cdot m$. Next we divide the width n into a number of segments equal or proportional to p, whereby each segment has a length proportional to $n/p = q$. At the points of division we rule the rectangle by parallel lines into strips, whose width is thus proportional to q.

Such "ruled profiles" exhibit, properly integrated, all the significant data entering into three-dimensional relations. Although likely to be less accurate in the presentation of the factors p and q, they have the advantage of being more easily read and interpreted than the more complex solid figures.

Figures 1 and 2 show examples of profiles of two-dimensional relations; Figure 3 of a solid, and Figure 4 of a ruled profile of three-dimensional relations.[6]

C. *Composites.*[7] The profile method appears to be of special advantage where the question is to represent not merely the data of a class (n, m, A) or (p, q, m, A) as a whole, but the distribution of these data over a number of component classes into which the whole class is divided.

---

[6] Among the many other data which can advantageously be presented as two-dimensional profiles we mention: volume of crops (=acreage ×bu/acre); earnings of certain classes of wage earners (=working time×wage rate); workers involved in labor board cases (=number of cases×workers×case); passenger-miles or ton-miles of traffic (=number of passengers or tons×average mileage). Of three-dimensional relations annual man-hours in an industry (=average employment×daily hours×average working days in the year); annual earnings (=daily hours×wage rate×average working days in the year.) Both these are of great interest in connection with the question of a guaranteed annual wage.

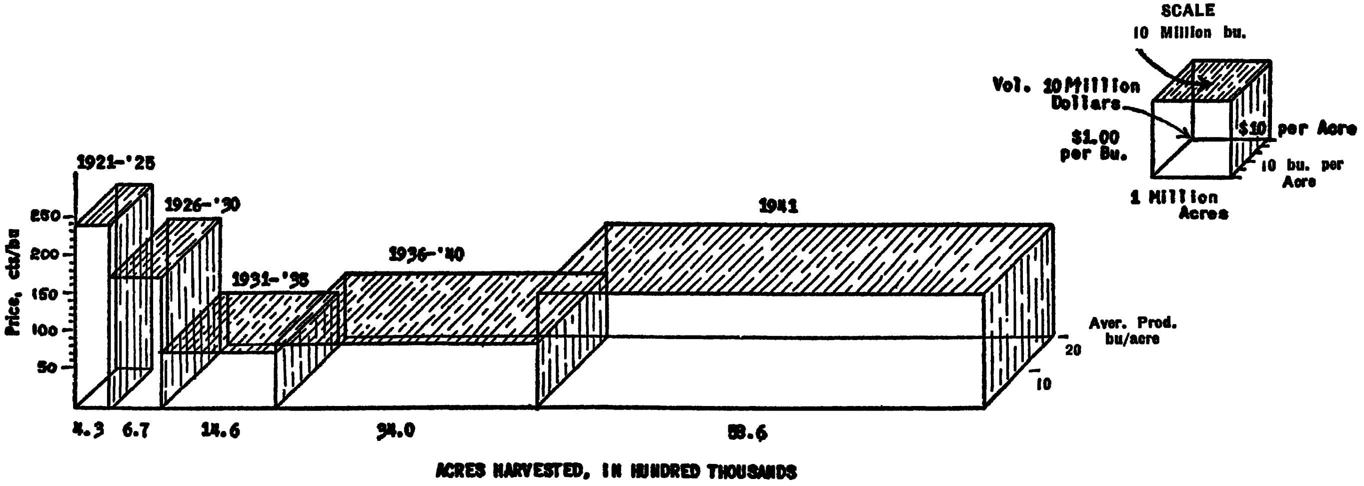
[7] The figures here described as "composites" are the only type with which Karsten (*op. cit.*) illustrates his "Vertical Area-Bar-Charts"; neither the two and three dimensional, nor the ruled time-series profiles are given or described.

FIGURE 3

SOY BEAN PRODUCTION IN THE UNITED STATES, 1921–1941

(1921–1940 five-year averages; 1941 single year)

Value = Acreage × Average Productivity × Price



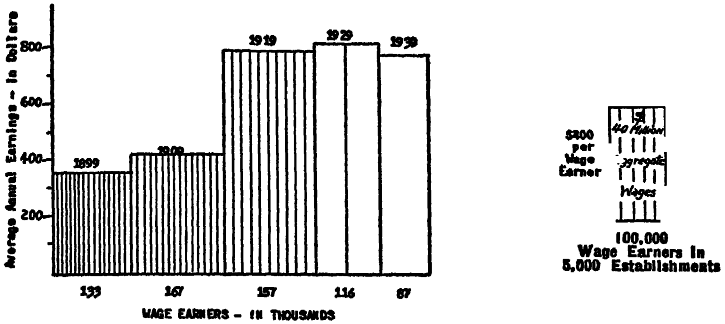
Source: U. S. Statistical Abstract, 1942, p. 776

Note: This profile integrates the simultaneous influences of the three separately varying dimensions: rapidly expanding acreage (width), and increasing productivity per acre (depth), together resulting in rapidly increasing total crops (shaded top faces); and abruptly fluctuating price (height), giving rise at times to smaller values per acre (shaded lateral faces), but steadily growing total values (volumes of solids).

FIGURE 4

EMPLOYMENT AND EARNINGS IN TOBACCO MANUFACTURING
UNITED STATES, 1899–1939

Aggregate Wages = Number of Wage Earners × Average Annual Earnings



Source: U. S. Census of Manufactures, 1939, II, 1, p. 270

Note: In this "ruled profile" several employment data pertaining to the tobacco industry are shown dimensionally integrated: total number of wage earners, (width), total wages (area); average annual earnings (height) number of establishments to nearest thousand (number of internal strips); average number of workers per establishment (width of strips). The profile thus shows at a glance the development from most poorly paid production in many very small establishments towards larger scale production in ever fewer establishments, employing in toto fewer workers at higher, but still very low average wages.

The customary presentation of a distribution of data over component classes is either by way of compound bars; or by circular charts showing the share of each component as sectors of the full circle; or, where the time variations of the several components are of principal interest, by way of compound line graphs showing the components added vertically onto each other. To present in such manner the triplets of data of a two-dimensional relationship distributed over *k* components one would need three separate *k*-fold compound line or bar graphs; or pairs of circular charts, for the presentation of the *n* and *A* data, combined with a bar graph for the *m* data. And at that these multiple compound graphs would fail to show the dimensional interrelationships between the particular triplets (n_i, m_i, A_i) belonging to one same component class.

Under the dimensional scheme one combines such triplets for each component class into one rectangle, and places all *k* such component rectangles side by side. The figure so obtained consisting of *k* component rectangles all of which belong to one common time interval, we call "composites," (in contrast to the "profiles" which consist of rectangles belonging to successive time intervals).

In composites the disadvantages of the dimensional presentation are far smaller than in the profiles, because they concern a distribution of values at a given time only, not their variation in time. On the other hand the advantage of showing all *k* triplets of data in their proper dimensional relationship in a single figure is very considerable.

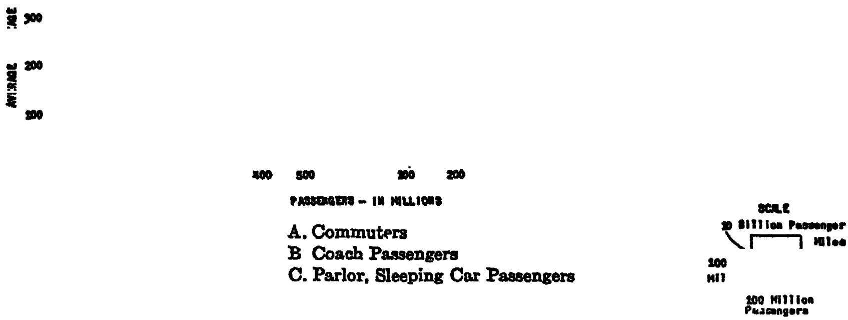
Finally one can construct "composite profiles," that is, sequences of composites for successive time intervals, which are capable of giving a very compactly informative view of the changing aspects of complex phenomena.[8]

The three figures below show illustrations of various types of composites: Figure 5 of a simple distribution of two-dimensional data; Figure 6 of a special kind, which we call "frequency composite"; Figure 7 of a sequence of "ruled" composites.[9]

FIGURE 5

RAILROAD TRAVEL, BY CLASSES OF PASSENGERS
UNITED STATES, 1941, 1943

Passenger-Miles = Passengers × Average Mileage



Source: Statistics of Railways in the U. S. 1943, Tables 44, 52

Note: This figure shows data for railroad traffic: number of passengers (width), average mileage (height), and passenger-miles (area) distributed over three classes of passengers. The several influences of wartime conditions are seen, dimensionally integrated: the increased number but constant mileage of the commuters; and the increase, both in number and mileage of actual travelers.

---

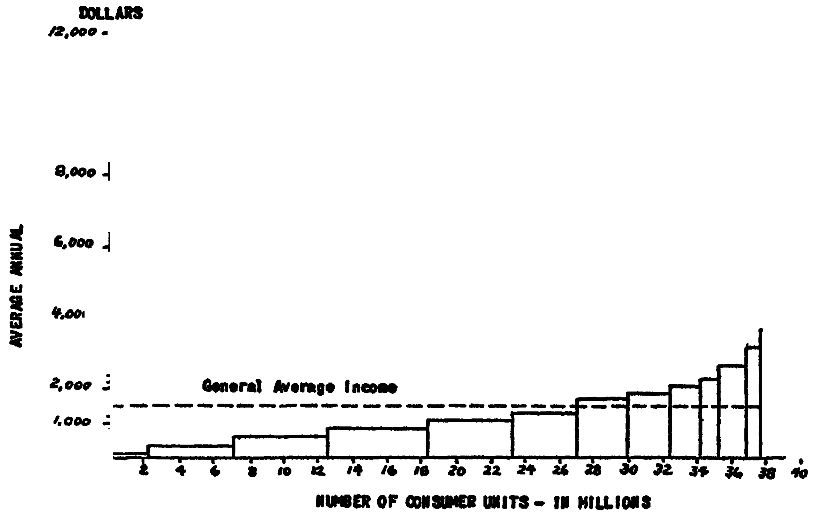
[8] See "Strike Profiles," *loc. cit*, figs. 2, 3, 4, 5.

[9] Other data which can effectively be presented as composites are: volumes of crops (acreage ×bushel/acre) distributed over various geographical regions, states, or countries; values of crops (acreage×value per acre) distributed over various crops; payrolls (workers×average pay) distributed over different industries, or different classes of personnel; traffic data, as in Figure 5, distributed over various means of transportation (railroads, airplanes), and many others.

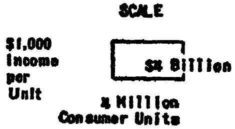
FIGURE 6

CONSUMERS. INCOME, BY INCOME LEVELS, UNITED STATES 1935

Aggregate Income = Consumer Units × Average Income



Source: U. S. Statistical Abstract, 1942, Table 384.

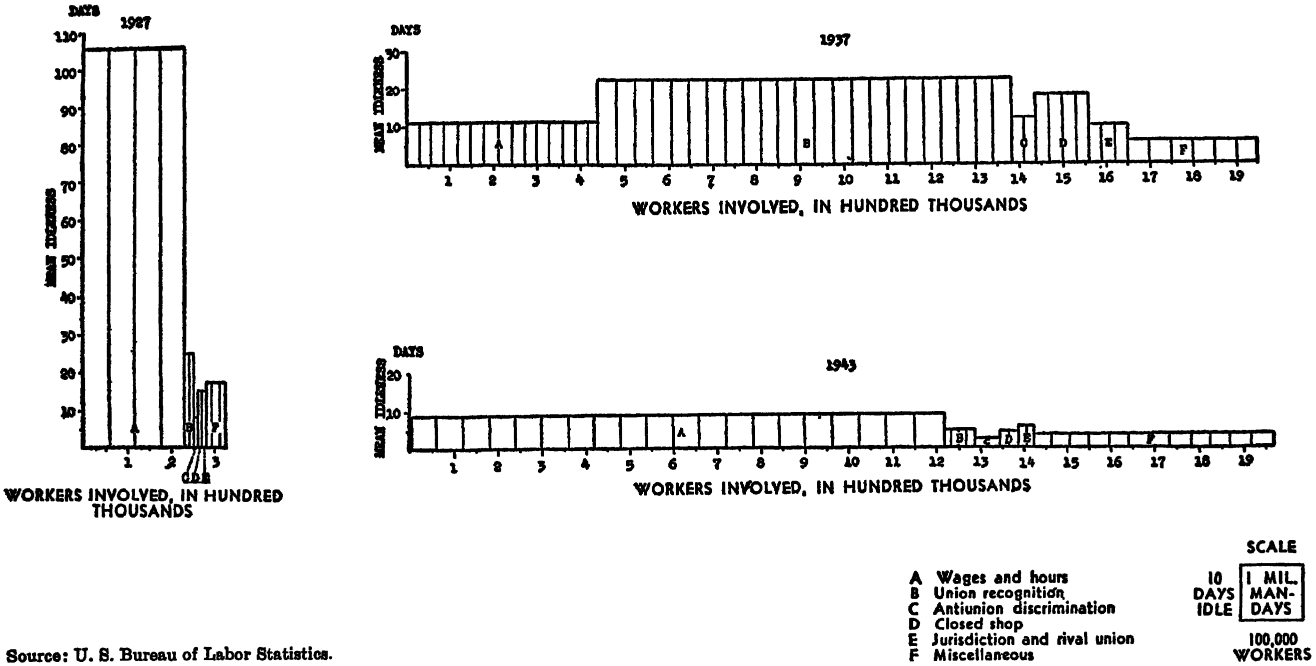


Note: The special feature in this "frequency composite" is that the component classes (income levels) are simply magnitude classes of the dimension $\bar{m}$ (average income) itself. Consequently the data constitute essentially a frequency distribution, the width (n_i) showing how many elements have $\bar{m}$-values in the interval $(m_{i+1} - m_{i-1})$, (the heights of the two adjacent rectangles). However this composite shows additionally also for each income level the aggregate income (area A_i), and the average income (height $\bar{m}_i$), and thus gives more information than ordinary frequency graphs, simple and cumulative.

D. *Historical Profiles.* Many apparently one-dimensional line or bar graphs of time series are really two-dimensional profiles, although of a different kind than those discussed above.

FIG. 7. STRIKES BY ISSUES, UNITED STATES, SELECTED YEARS

Man-Days Idle = Workers Involved × Mean Duration of Idleness

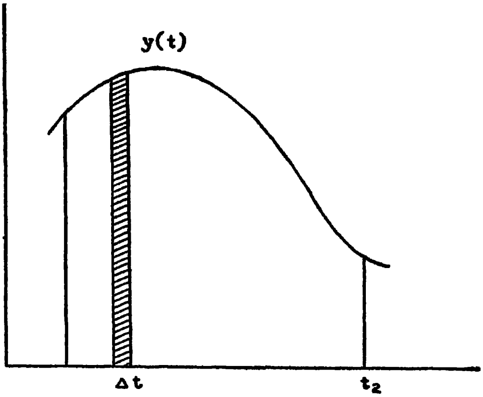


Source: U. S. Bureau of Labor Statistics.

Taken from fig. 5 of "Strike Profiles," *loc. cit.*

Note: These composites show annual strike data classified by strike issues. Each rectangle represents the number of workers involved (width), mean duration of idleness (height), man-days idle (area) in strikes over the several principal issues (indicated by letter symbol). In addition the number of internal strips shows the number of strikes to the nearest hundred; hence the width of the strips gives the average number of workers per strike. One thus obtains a comprehensive, yet numerically detailed view of the characteristic differences in the strike situation in the three selected years, that is, of the varying prominence of the several issues as to workers involved, time lost, and number and average magnitude of strikes.

FIGURE 8



As an example let us take a line graph showing as its ordinate y a labor force varying in time t plotted along the horizontal time axis, (Fig. 8). Let Δt, in days, be a time interval during which y is sufficiently constant. Then the product $y \cdot \Delta t = \Delta A$, which is the strip of area under the curve standing over Δt, gives the number of man-days of employment. Hence the total area under the curve between two given dates t_1 and t_2,

$$A = \sum_{t_1}^{t_2} y \cdot \Delta t$$

gives the total of man-days of employment during the time $t_2 - t_1$. The figure, therefore, is a profile showing three dimensionally related data: the number y of workers employed at any time t; the duration Δt of employment of y men, and the man-days of employment A.

In other instances the height y of a line or bar graph represents, not as in the above case, a magnitude which *endures in time*, but a number of occurrences taking place and accumulating per time unit, that is, a time rate, $\Delta N / \Delta t$ (e.g. the number of births per year; or the number of strikes per month; or the number of workers quitting their jobs per month, or per week, etc.). In such cases the area under the curve

$$A = \sum_{t_1}^{t_2} y \cdot \Delta t = \sum_{t_1}^{t_2} \frac{\Delta N}{\Delta t} \cdot \Delta t = \sum_{t_1}^{t_2} \Delta N = N$$

represents the total number of occurrences during the period $t_2 - t_1$.

Hence again the graph is truly a profile representing three dimensionally related data: the time rate y, the time Δt over which the rate prevails, and the total number of occurrences during the time interval in question.[10]

Profiles of this kind, in which chronological time is represented by a time axis, we call "historical profiles," in contrast to the statistical profiles discussed above.

A special form of historical profile results when the phenomenon represented is intermittent rather than continuous, that is, occurs at certain times, lasts for a while, and ceases again. The profile then consists of separate, quasi-rectangular figures along the time axis, their place showing the time of occurrence, their height the "magnitude," their width the duration, and their area the "time integral" of each particular occurrence.

Profiles of this kind are useful when individual occurrences rather than statistical summaries are important. For example, where individual strikes rather than overall strike statistics are of interest—as are, indeed, from the standpoint of management the strikes in its own establishments—"historical strike profiles" could be used to considerable advantage. Any one strike would be shown at its proper place along the time axis by plotting vertically the day-to-day number of workers involved, whereby the area would give the number of man-days idle. Such profiles show at a glance when, how often, and for how long how many men were out, and how many man-days were lost.[11]

On similar principles, only on a time scale ranging over centuries rather than months and years, one can construct very instructive "historical war profiles," for individual nations or groups of nations, by using for the height the absolute or relative "number of men involved," that is, the size of the armies, as a rough yardstick for the magnitude of wars (at least in the pre-atomic era), while the width of the rectangles shows the duration, and the areas the "man-years" of wars.

---

[10] These dimensional relations are familiar in the field of physics: Two examples analogous to the situations discussed above are: when a force F is plotted as a function of time t, the area under the curve, $\int_{t_1}^{t_2} F\,dt$ represents the Impulse; and when a velocity $v = ds/dt$ is plotted as a function of time t, the area under the curve $\int_{t_1}^{t_2} v\,dt$ represents the total distance $s_2 - s_1$ traversed during the time interval $t_2 - t_1$.

[11] To render such profiles more informative one could show further, say by distinctive coloring, the principal causes of the various strikes; through shaded extensions of the strike rectangles one could show the number of workers made indirectly idle through the strike. To allow the representation, in one same diagram, of large and of small strikes the vertical scale could be made logarithmic.

A METHOD FOR OBTAINING AND ANALYZING SENSITIVITY DATA*

W J Dixon
University of Oregon
AND
A M. Mood
Iowa State College

The standard method of dealing with sensitivity of dosage-mortality data is the probit technique developed by Bliss and Fisher. This paper provides an alternative technique based on a special system for obtaining such data It has some advantages when observations must be taken on individuals rather than groups of individuals, and it may be preferred in certain other situations

INTRODUCTION

EXPERIMENTAL investigations often deal with continuous variables which cannot be measured in practice For example, in testing the sensitivity of explosives to shock, a common procedure is to drop a weight on specimens of the same explosive mixture from various heights. There are heights at which some specimens will explode, and others will not, and it is assumed that those which will not explode would explode were the weight dropped from a sufficiently greater height. It is supposed, therefore, that there is a critical height associated with each specimen, and that the specimen will explode when the weight is dropped from a greater height and will not explode when the weight is dropped from a lesser height. The population of specimens is thus characterized by a continuous variable—the critical height—which cannot be measured. All one can do is select some height arbitrarily and determine whether the critical height for a given specimen is less than or greater than the selected height.

This situation arises in many fields of research. Thus in testing insecticides, a critical dose is associated with each insect, but one cannot measure it. He can only try some dose and observe whether or not the insect is killed, that is, observe whether the critical dose for that insect is less than or greater than the chosen dose. The same difficulty arises in pharmaceutical research dealing with germicides, anesthetics, and other drugs; in testing strength of materials; in psycho-physical research dealing with threshold stimuli; and in several areas of biological and medical research.

---

* This paper is in part an adaptation of a memorandum submitted to the Applied Mathematics Panel by the Statistical Research Group Princeton University The Statistical Research Group operated under a contract with the Office of Scientific Research and Development, and was directed by the Applied Mathematics Panel of the National Defense Research Committee.

In true sensitivity experiments it is not possible to make more than one observation on a given specimen. Once a test has been made the specimen is altered (the explosive is packed; the insect is weakened) so that a bona fide result cannot be obtained from a second test. The common procedure in experiments of this kind is to divide the sample of specimens into several groups (usually but not necessarily of the same size) and to test one group at a chosen level, a second group at a second level, and so on. The data consist of the numbers affected and not affected at each level. A method of analyzing such data (variously called "sensitivity" data, "all or none" data, "quantal responses") has been developed by Bliss and Fisher [references 1, 2], and discussed by other writers [3, 4, 5, 6].

A new technique for obtaining sensitivity data has been developed and used in explosives research. The authors became acquainted with this new method in 1943 at the Explosives Research Laboratory, Bruceton, Pennsylvania. It has come to be called the "up and down" method. The method may be employed in any sensitivity experiment, but we shall discuss it in terms of the explosives to avoid general terminology.

The technique is to choose some initial height h_0, and a succession of heights $h_1, h_2, h_3, \cdots$ above h_0 together with a succession $h_{-1}, h_{-2}, h_{-3}, \cdots$ below h_0. The first specimen is tested by dropping the weight from height h_0. If the first specimen explodes, the second specimen will be tested at h_{-1}, otherwise the second specimen will be tested at h_1. In general, any specimen will be tested at the level immediately below or immediately above the level of the previous test according as there was or was not an explosion on the previous test. The result of such an experiment might be portrayed as in Figure 1 where the x's represent explosions and the o's non-explosions. The first test is on the left at the highest level; this was a success (explosion) so the second test was made at the next lower level and was also a success; the third test was therefore made at the level below that of the second and since it was a failure the fourth test was made at the level above that of the third test.

The primary advantage of this method is that it automatically concentrates testing near the mean. We shall see later that this increases the accuracy with which the mean can be estimated. Or in other words,

RECORD OF A SAMPLE OF SIXTY TESTS

Normalized Height		Number of x's	Number of o's
2 0	x	1	
1 7	x x x x x x x x x x	10	
1 4	o x o x x x x x o o x o x x x o x x x o x x o x x o x	18	9
1 1	o x o o o o o o o o o o o o o o o x o	2	18
0 8	o o		2

FIGURE 1

for a given accuracy the up and down method will require fewer tests than the ordinary method of testing groups of equal size at preassigned heights. The saving in the number of observations may be of the order of 30 to 40 per cent (see Appendix A).

Another advantage is that the statistical analysis is quite simple in certain circumstances whereas the analysis for the ordinary method is rather tedious.

The method has one obvious disadvantage in certain kinds of experiments because it requires that each specimen be tested separately. This is not important in explosives experiments because each test must be made separately anyway. But in tests of insecticides, for example, a large group of insects can sometimes be treated as easily as a single one, and in large experiments of this kind any advantage of the up and down method might well be outweighed by this requirement of single tests. Even here, if expensive laboratory animals were being used, the advantage in economy of tests might offset the trouble of making single tests.

CONDITIONS ON THE EXPERIMENT

The statistical analysis of data obtained can be quite simple provided the experiment satisfies certain conditions. Less restrictive conditions must be fulfilled in order that any analysis will be possible. These will be discussed here and the actual analysis will be given in the following section.

In the first place, the analysis requires that the variate under analysis be normally distributed. In practice the variate of interest to the research worker can rarely be considered to be normally distributed. It is therefore necessary that the natural variate be transformed to one which does have the normal distribution. This is readily done provided the research worker has enough experience and data on his material to be able to specify rather accurately the shape of his distribution function. It is often the case in dosage mortality experiments and in experiments on explosives that the logarithm of the dosage concentration or of the height is reasonably normally distributed. But in other areas of research, and sometimes in these areas, other transformations are more appropriate [7].

If one has no idea of the shape of his distribution function then the data of the experiment itself must be used to provide this information. The common procedure here is to compute the percentage affected at each level and plot these percentages on arithmetic probability paper against various functions of the variate in question. Usually one can soon discover what sort of function will force the percentages to lie sensibly along a straight line. There are, of course, infinitely many functions to choose from; the chosen function should be as simple as possible consistent with whatever knowledge is available concerning the nature of the material at hand.

We have already mentioned that the up and down method is particularly effective for estimating the mean. It is not a good method for estimating small or large percentage points (for example, the height at which 99 per cent of specimens explode) unless normality of the distribution is assured. In fact no method which uses the normal distribution can be relied on to estimate extreme percentage points because such estimates depend critically on the assumption of normality. In most experimental research, it is possible to find simple transformations which make the variate essentially normal in the region of the mean, but to make it normal in the tails is quite another matter. Nothing short of an extensive exploration of the distribution involving perhaps thousands of observations will suffice here. Bartlett [8] has recently presented an interesting technique for dealing with this problem.

A second condition on the experiment is that the sample size must be large if the analysis to be described is to be applicable. As it turns out, the effective sample size is only about half the actual sample size. The statistical analysis is based on large sample theory so that if one uses the analysis on a sample of size forty, he will in effect be using large sample theory on a sample of size twenty. Measures of reliability may well be very misleading if the sample size is less than forty or fifty.

A further condition is necessary if the statistical analysis is to be simple. One must be able to estimate roughly in advance the standard deviation of the normally distributed transformed variate. The interval between testing levels should be approximately equal to the standard deviation. This condition will be well enough satisfied if the interval actually used is less than twice the standard deviation. This requirement is not severe, for research workers who repeatedly perform these requirements on essentially similar materials can usually make very good preliminary estimates. This is the case in explosives research or biological assay, for example. This circumstance (of repeated experiments) is precisely the one in which a simple analysis is most desirable.

STATISTICAL ANALYSIS

The simple method of analysis given in this section is applicable only when all the conditions described in the preceding section are fulfilled. The theory underlying the method is given in Appendix A. The more complex analysis required when the levels are not equally spaced or when the distance between levels exceeds twice the standard deviation is given in Appendix B.

We again revert to the explosives experiment in describing the method. Suppose it is known for the given type of explosive that the logarithms of the critical heights are normally distributed. Letting h represent the height, $y = \log h$ will then be the normally distributed variate. We shall call y the normalized height, and represent the mean and variance of its distribution by μ and σ^2. The experiment is performed by choosing an initial height for the first test, say h_0. This should be chosen near the anticipated mean. The other testing levels are determined so that the values of the normalized height y are equally spaced. If d is the preliminary estimate of σ, and if $y_0 = \log h_0$, then the actual testing heights are obtained by putting $\log h = y_0 \pm d$, $y_0 \pm 2d$, $y_0 \pm 3d, \cdots$, and solving for h. The heights will then be so spaced that the transformed variate is equally spaced with spacing equal to its anticipated standard deviation. All computations are done in terms of y.

In any experiment the total number of successes will be approximately equal to the total number of failures. In fact, the number of failures at any level cannot differ by more than one from the number of successes at the next higher level. For estimating μ and σ only the successes or only the failures are used, depending on which has the smaller total. In the example shown in Figure 1 there are fewer failures than successes so the failures would be used. We shall let N denote the smaller total and let $n_0, n_1, n_2, \cdots n_k$ denote the frequencies at each level for this less frequent event where n_0 corresponds to the lowest level and n_k the highest level on which the event occurs. We have then $\Sigma n_i = N$.

The estimates of μ and σ are based on the first two moments of the y values using the frequencies n_i. But since the y values are equally spaced, the moments are more easily computed in terms of the two sums

$$A = \sum in_i$$
$$B = \sum i^2 n_i.$$

In this notation, the estimate of μ, say m, is

$$m = y' + d\left(\frac{A}{N} \pm \frac{1}{2}\right) \tag{1}$$

where y' is the normalized height corresponding to the lowest level on which the less frequent event occurs. The plus sign is used when the analysis is based on the failures, and the minus sign when it is based on the successes.

The sample standard deviation is

$$s = 1.620d\left(\frac{NB - A^2}{N^2} + .029\right) \tag{2}$$

and this, of course, is the estimate of σ. This is a curious estimate in that while it is a linear function of $(NB-A^2)/N^2$, it gives the estimate of the standard deviation, not the square of the standard deviation. The formula is an approximate one which is quite accurate when $(NB-A^2)/N^2$ is larger than 0.3 but breaks down rapidly when $(NB-A^2)/N^2$ becomes less than 0.3. In the latter instance the formula cannot be used and the more elaborate calculation described in Appendix B must be employed.

The example of Figure 1 will illustrate the use of the formulas. Here the y values used were 2, 1.7, 1.4, 1.1, 0.8; the level of the first test y_0 being 2, and d being 0.3. Among the sixty tests there were 31 explosions and 29 failures, hence the latter are used to estimate the parameters. The failures appear on three levels (0.8, 1.1, 1.4) with frequencies $n_0 = 2$, $n_1 = 18$, $n_2 = 9$. We have then $N = 29$, $A = 36$, $B = 54$, so that the mean is

$$m = 0.8 + 0.3\left(\frac{36}{29} + \frac{1}{2}\right) = 1.32$$

and the standard deviation is

$$(1.620)(0.3)\left(\frac{270}{841} + .029\right) = .17.$$

The sample was actually drawn from a normal population with $\mu = 1.3$ and $\sigma = 0.2$ using Mahalanobis' [9] table of random normal deviates. The mean and standard deviation of the sixty observations were 1.312 and .158 so that it was a fairly representative sample.

Percentage points would be estimated by $m + ks$ where k is chosen from tables of the normal deviate to give the desired percentage. Thus in the example, the 5 per cent point is estimated by

$$1.32 - (1.645)(.17) = 1.04.$$

If the y values are thought of as logarithms to base ten of actual heights in inches in an explosives experiment, the antilogarithms of estimated percentage points would be estimates of the corresponding points for the distribution of h. Thus the median (not mean) value of h is estimated by

$$\text{antilog } 1.32 = 20.9 \text{ inches}$$

and the 5 per cent height by

$$\text{antilog } 1.04 = 11 \text{ inches.}$$

The antilogarithm of s does not estimate the standard deviation for h, however, and any computation which involves the standard deviation (estimates of percentage points, confidence limits) must be done in terms of the normalized height, and only the final result transformed to actual heights.

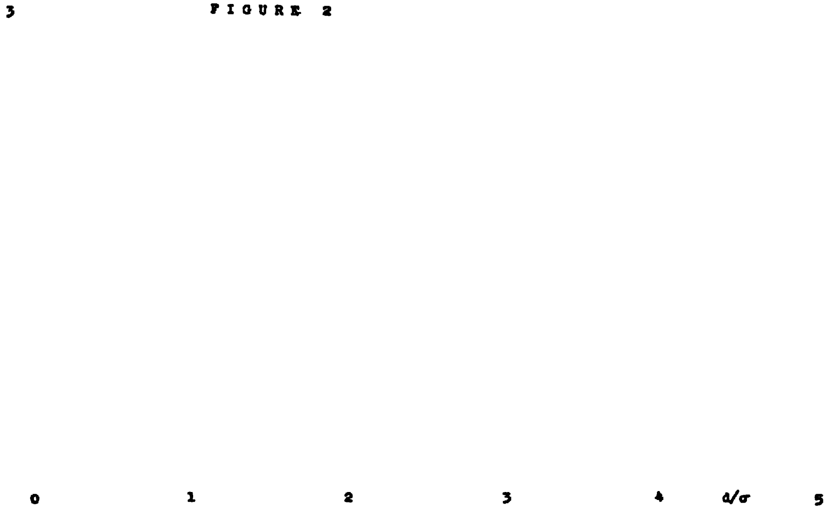
CONFIDENCE INTERVALS

Ordinarily the standard deviation of a sample mean, m, is given by $\sigma_m = \sigma/\sqrt{N}$ where σ is the population standard deviation and N the sample size. In the present case this expression must be multiplied by a factor which we shall call G, so that the formula for the standard error of the mean is

$$\sigma_m = G\sigma/\sqrt{N} \tag{3}$$

and G depends on the ratio d/σ and on the position of the mean relative to the testing levels. G is plotted in Figure 2 as a function of d/σ. The position of the mean relative to the testing levels does not affect G unless the interval d is large; the solid branch of the curve gives the value of G when the mean falls on one of the testing levels, while the dashed branch gives the value when the mean falls midway between two levels. Curves for other positions of the mean would fall between the two branches.

FIGURE 2



In practice σ is not known and s must be used in (3) to obtain an estimate say, s_m, of σ_m. In the illustrative example with $s=.17$, we have $d/s=1.8$ so that G is about 1.12. The estimate of σ_m is therefore

$$s_m = (.17)(1.12)/\sqrt{29} = .035.$$

A confidence interval for m may now be estimated by $m \pm ks_m$. Thus a 95 per cent confidence interval is

$$1.32 \pm (1.96)(.035) \text{ or } 1.25 \text{ to } 1.39$$

using large sample theory. For moderate values of N, it might be preferable to use the value of k given by the t distribution for $N-1$ degrees of freedom, but it is likely that this is a minor matter relative to the error of using large sample theory for moderate values of N. Again assuming the confidence interval refers to the logarithm of an actual height, it gives rise to an asymmetric 95 per cent confidence interval (18 to 25 inches) for the median height.

The standard error of the sample standard deviation, say σ_s, is ordinarily given by $\sigma/\sqrt{2N}$, but in the present analysis an additional factor is again required. We shall write

$$\sigma_s = H\sigma/\sqrt{N} \tag{4}$$

where we have incorporated the $1/\sqrt{2}$ into the extra factor. H is plotted in Figure 2 where the solid branch gives the value of H when the mean falls on a level, while the dashed branch gives the value when the mean is midway between two levels. When d/σ is less than two there will be little error introduced by interpolating linearly between the two branches for other positions of the mean. Thus if the mean falls $d/4$ from a testing level, one may use the value of G midway between the two branches. For the illustrative example with $d/s=1.8$, we find H to be about 1.24 so that the estimate of σ_s is

$$s_s = (1.24)(.17)/\sqrt{29} = .039.$$

The estimate s_s would be used to estimate the standard error of a percentage point, $m+ks$; the estimate would be $\sqrt{s_m{}^2+k^2s_s{}^2}$. Thus in the example, a 95 per cent confidence interval for the 5 per cent point would be estimated by $1.04\pm(1.96)\sqrt{(0\ 35)^2+(1\ 645)^2(.039)^2}$ or 88 to 1.20. We should mention again that the estimation of small or large percentage points depends strongly on the assumption of normality in the tails. It can easily happen that a relatively small error in this assumption may far outweigh the sampling error indicated by the confidence interval, especially in the case of very extreme percentages, say 1 per cent or 0.1 per cent.

CHOICE OF TESTING INTERVAL

The curves in Figure 2 have been extended beyond $d=2\sigma$ in order to show what happens to the measures of precision for larger intervals. Curve G shows that the precision of the mean steadily decreases as d increases. The two branches of H show that there is an optimum spacing for estimating the standard deviation depending on the position of the mean relative to the testing levels. Since the mean is usually unknown, this information is of little practical value.

Curve G indicates that the interval should be quite small for maximum precision in the mean, but in practice this is not true for several reasons. In the first place the curves are for expected values and essentially assume infinite sample sizes, and in fact very large samples are required to get good estimates of the mean for a very small interval. The estimate may be biased appreciably toward the initial testing level unless the sample is very large. Secondly, a small interval may cause one to waste observations unless a good choice for the initial level is made. If a poor choice is made, many observations must be spent getting from that level to the region of the mean. And finally, since σ is usually unknown, the precision of the mean must actually be measured by s, and the accuracy of s becomes poor for very small intervals as shown by curve H.

All these considerations indicate that the interval should be within the range of about 0.5σ to 2σ, and experiments with the method support this conclusion.

APPENDIX A

If y is normally distributed with mean μ and variance σ^2, and if tests are made at

$$y_i = y_0 \pm id \qquad i = 0, 1, 2, \cdots \tag{5}$$

where y_0 is the level of the initial test, then there will be, say, n_i successes and m failures at y_i, and the distribution of these latter variates is

$$P(n, m | y_0) = K \prod_{i=-\infty}^{\infty} p_i^{n_i} q_i^{m_i} \tag{6}$$

where

$$p_i = \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1/2(t-\mu)^2}{\sigma^2}} dt = 1 - q_i \tag{7}$$

and where K is not a function of μ and σ^2.

The estimation of μ and σ^2 is based on the principle of maximum likelihood. We shall not maximize (6) directly, however, because a material simplification in the analysis can be made by neglecting a small part of the information in the sample. It is clear that

$$| n_i - m_{i-1} | = 0 \quad \text{or} \quad 1$$

so that either one of the sets (n_i) or (m_i) contain practically all the information in the sample. If $N = \Sigma n_i$ and $M = \Sigma m_i$, and assuming $N \leqq M$, we may write (2) in the form

$$P(n, m | y_0, M - N) = K' \prod_i (p_i q_{i-1})^{n_i} \tag{8}$$

and this is the expression which will be maximized. Even if $M-N$ is not small, only a small amount of information is being neglected, because in this instance the initial level will have been poorly chosen and these neglected observations will have been spent in getting from y_0 to the region of the mean; they will obviously contribute little to the more precise location of the mean.

On putting the derivatives of (8) with respect to μ and σ equal to zero we have the relations

$$\sum n_i \left(\frac{z_{i-1}}{q_{i-1}} - \frac{z_i}{p_i} \right) = 0 \tag{9}$$

$$\sum n_i \left(\frac{x_{i-1} z_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i} \right) = 0 \tag{10}$$

where z_i represents the ordinate of the distribution of y at y_i and $x_i = (y_i - \mu)/\sigma$. The expected values of the left hand sides of these two expressions are readily found to be zero on substituting $E(n_i)$ for n_i. $E(n_i)$ may be determined from the relation

$$\frac{E(n_{i+1})}{q_i} = \frac{E(n_i)}{p_i}. \tag{11}$$

If we let

$$\begin{aligned} w_0 &= 1 \\ w_i &= \prod_{j=0}^{i-1} \frac{q_j}{p_j} \qquad i > 0 \\ &= \prod_{j=-1}^{i} \frac{p_j}{q_j} \qquad i < 0 \end{aligned}$$

then it follows that

$$E(n_i) = N w_i \Big/ \sum_{-\infty}^{\infty} w_i. \tag{12}$$

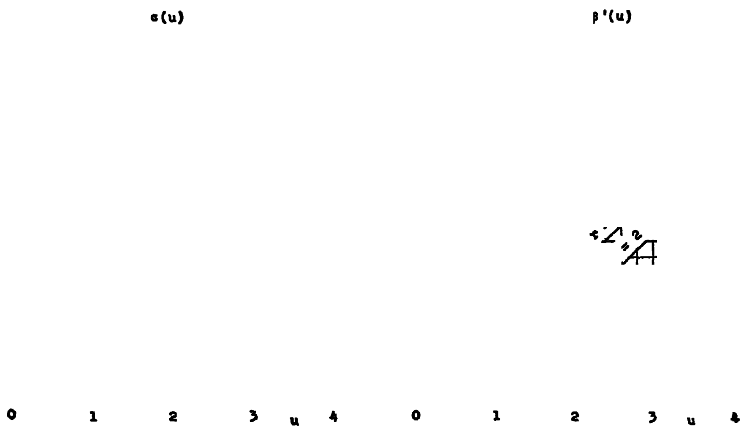
The maximum likelihood estimates of μ and σ are the roots, say $\hat{\mu}$ and $\hat{\sigma}$, of equations (9) and (10). While there is no simple closed expression for these roots, it turns out that they can be very closely approximated when $d < 2\sigma$. The function

$$\alpha(u) = \frac{z(x)}{q(x)} - \frac{z(x + d/\sigma)}{p(x + d/\sigma)}$$

where $u = x + d/2\sigma$, is nearly linear in u when $d < 2\sigma$. This is illustrated in Figure 3. Similarly

$$\beta(u) = \frac{xz(x)}{q(x)} - \frac{(x + d/\sigma) z(x + d/\sigma)}{p(x + d/\sigma)}$$

is nearly quadratic in u as is indicated by the graph of its first derivative in Figure 3, where $r = d/\sigma$. We may conclude therefore that the



estimates are essentially determined by the first two moments of the y_i using the n_i as weights.

If we let

$$\hat{\mu}_1 = \frac{1}{N} \sum n_i y_i, \qquad \hat{\mu}_2 = \frac{1}{N} \sum n_i y_i^2 \tag{13}$$

we find

$$E(\hat{\mu}_1) = \mu - d/2 \tag{14}$$

and

$$E(\hat{\mu}_2) - E^2(\hat{\mu}_1) + d^2/4 = \sigma^2 \frac{\sum w_i x_i^2}{\sum w_i} \cdot \tag{15}$$

The expression on the right of (15) is nearly linear in σ when $d<2\sigma$, and its linear approximation was used to determine the estimate of σ given in equation (2). The function is plotted in Figure 4; the solid branch represents the function when the mean falls at one of the y_i, and the dashed branch when the mean falls midway between two levels. The two branches diverge rapidly as d becomes larger than 2σ.

The variance and covariances of $\hat{\mu}$ and $\hat{\sigma}$ are determined from the second derivatives of $L = \log P$ where P is defined by (8). The expected values of the derivatives are readily found to be

$$E\left(\frac{\delta^2 L}{\delta\mu^2}\right) = -\frac{N}{\sigma^2}\sum w_i\left(\frac{z^2{}_{i-1}}{q^2{}_{i-1}} + \frac{x_i{}^2}{p_i{}^2}\right)\Big/\sum w_i = -\frac{N}{\sigma^2 G^2} \qquad (16)$$

$$E\left(\frac{\delta^2 L}{\delta\mu\delta\sigma}\right) = \frac{N}{\sigma^2}\sum w_i\left(\frac{x_{i-1}z^2{}_{i-1}}{q^2{}_{i-1}} + \frac{x_i z_i{}^2}{p_i{}^2}\right)\Big/\sum w_i \qquad (17)$$

$$E\left(\frac{\delta^2 L}{\delta\sigma^2}\right) = -\frac{N}{\sigma^2}\sum w_i\left(\frac{x^2{}_{i-1}z^2{}_{i-1}}{q^2{}_{i-1}} + \frac{x_i{}^2 z_i{}^2}{p_i{}^2}\right)\Big/\sum w_i$$

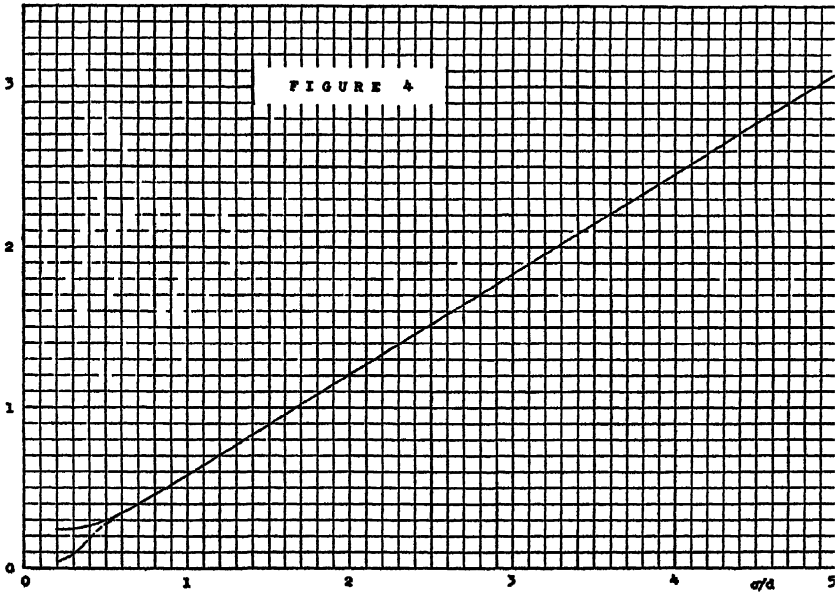
$$N$$

$$\sigma^2 H^2 \qquad (18)$$

Expression (17) does not vanish unless the mean falls on a level or midway between two levels. However, we have regarded the covariance as being negligible for all practical purposes. It gives rise to a maximum correlation between $\hat{\mu}$ and $\hat{\sigma}$ of the order of .0002 when $d = \sigma$, and .02 when $d = 2\sigma$. We have then

$$\sigma_\mu{}^2 = G^2\sigma^2/N, \qquad \sigma_\sigma{}^2 = H^2\sigma^2/N \qquad (19)$$

approximately, where G and H are defined in (16) and (18). These are the functions plotted in Figure 2.



It is not possible to make a very satisfactory comparison between this method and the ordinary probit method, but the following computations provide some indication of the relative efficiencies. Suppose $2N$ individuals are divided into five equal groups and tested at $y=0$, $\pm\sigma$, $\pm 2\sigma$. Bartlett [8], for example, shows that the variance of $\hat{\mu}$ for the probit analysis is about $5(.564)\sigma^2/2N$, whereas for the up and down method the variance is σ^2/N which when divided by the former value gives 71 per cent. When the sample is tested in six equal groups at $y=\pm\frac{1}{2}\sigma$, $\pm\frac{3}{2}\sigma$, $\pm\frac{5}{2}\sigma$, the ratio becomes 58 per cent. But these comparisons are not fair unless there is considerable uncertainty as to the general location of the mean. If the mean can be located to within, say, σ of its true position in advance of the experiment, then the efficiency of the probit method can be much improved by using groups of unequal size and testing the larger groups at levels thought to be near the mean.

APPENDIX B

When the chosen testing interval is larger than 2σ, or when the intervals are of unequal size, it is necessary to solve equations (9) and (10) for μ and σ. The intervals will be of unequal size, for example, when the normalizing transformation is unknown in advance of the experiment and must be deduced from the results of the experiment itself. A method of trial and error is probably as good as any other for solving the equations. One would first choose preliminary estimates, say m and s, of the roots by using equations (1) and (2) or simply by using guessed values. These preliminary estimates would be adjusted until the equations were satisfied to the desired degree of approximation. The left side of (9) will be positive when the trial value of μ is too small, and negative when it is too large. The left side of (10) will be positive when $s<\hat{\sigma}$, and negative when $s>\hat{\sigma}$. The equation (9) is relatively insensitive to changes in s, while the same is true of (10) for changes in m.

In order to facilitate the computations, the accompanying tables of z/p (Table I) and z/q (Table II) are provided. For negative values of x, p and q are interchanged, that is

$$\begin{array}{cc} z(x) & z(-x) \\ p(x) & q(-x) \end{array}$$

We shall illustrate the computation using the data of Figure 5. The normalized heights are .1, .9, 1.5, 1.9 as indicated in the Figure. We shall number the levels 0, 1, 2, 3 beginning with the lowest level. Since there are more successes than failures, the latter are used to determine the estimates. A preliminary estimate of μ may be obtained by using the average of the midpoints of the intervals weighted by the numbers n_i; thus we shall put

$$m_1 = 1/29[2(1.7) + 26(1.2) + (.5)]$$
$$\cong 1.2.$$

A rough estimate of σ may be determined by observing that the interval 0.9 to 1.5 appears to contain 26/29 or about 90 per cent of the distribution, hence we may use

$$1.645s_1 \cong \tfrac{1}{2}(1.5 - 0.9) = 0.3$$
$$s_1 = 0.18.$$

RECORD OF A SAMPLE OF SIXTY TESTS

Normalized Height		Number of x's	Number of o's
1.9	x x x	3	
1.5	x x x x x x x x x x o x x x x x x x x x x x x x x o x x x	27	2
.9	o o o o o o o o o o o o o o o x o o o o o o o o o o o	1	26
.1	o		1

FIGURE 5

In adjusting these estimates one might be tempted to adjust m_1 first by equation (9) and then go to equation (10) and adjust s_1 by using a good estimate of μ. It turns out, however, that the job can be done much more rapidly by considering both equations together. The following computational form may be used:

i	n_i	h_i	x_i	$n_i\left(\frac{z_{i-1}}{q_{i-1}}\ \frac{z_i}{p_i}\right)$	$\frac{x_i z_i}{q_i}$	$\frac{x_i z_i}{p_i}$	$n_i\left(\frac{x_{i-1}z_{i-1}}{q_{i-1}}\ \frac{x_i z_i}{p_i}\right)$
3	2	1.9	3.89	4.17		0	6.96
2	26	1.5	1.67	.00	3.48	.174	−9.05
1	1	.9	−1.67	−2.08	−.174	−3.48	3.48
0		.1	−6.11		0		
				2.09			1.39

Note that the table is arranged so that the frequencies of either the zeros or x's will be entered in the table as though they were x's. The symbol x_i represents $(h_i - m_1)/s_1$ where h_i is the height and m_1 and s_1 are the first approximations to μ and σ. The other computations are defined by the column headings. Thus the figure 4.17 at the top of the fifth column is obtained as $2(2.084 - .000)$; 2.084 being read from Table II at $x = 1.67$, and .000 being the value of z/p at $x = 3.89$ as shown by Table I. The sums, 2.09 and 1.39, of the fifth and eighth columns give the values of the left hand sides of equations (9) and (10) respectively; since both sums are positive. We conclude that both m_1 and s_1 are too small. Using $m_2=1.3$ and $s_2=.19$ we repeat the above calculation:

i	n_i	h_i	x_i	$n_i\left(\frac{z_{i-1}}{q_{i-1}}-\frac{z_i}{p_i}\right)$	$\frac{x_i z_i}{q_i}$	$\frac{x_i z_i}{p_i}$	$n_i\left(\frac{x_{i-1}z_{i-1}}{q_{i-1}}-\frac{x_i z_i}{p_i}\right)$
3	2	1.9	3.16	3.12		.01	3.28
3	26	1.5	1.05	−5.85	1.64	.282	−9.75
1	1	.9	−2.11	−2.47	−.093	−5.21	5.21
0		.1	−6.32		0		
				5.20			−1.26

TABLE I

VALUES OF z/p

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.798	0.792	0.785	0.779	0.773	0.766	0.760	0.754	0.748	0.741
0.1	0.735	0.729	0.723	0.717	0.711	0.705	0.699	0.693	0.687	0.681
0.2	0.675	0.669	0.663	0.657	0.652	0.646	0.640	0.634	0.629	0.623
0.3	0.617	0.612	0.606	0.600	0.595	0.589	0.584	0.578	0.573	0.567
0.4	0.562	0.556	0.551	0.546	0.540	0.535	0.530	0.525	0.519	0.514
0.5	0.509	0.504	0.499	0.494	0.489	0.484	0.479	0.474	0.469	0.464
0.6	0.459	0.454	0.449	0.445	0.440	0.435	0.430	0.426	0.421	0.417
0.7	0.412	0.407	0.403	0.398	0.394	0.389	0.385	0.381	0.376	0.372
0.8	0.368	0.363	0.359	0.355	0.351	0.346	0.342	0.338	0.334	0.330
0.9	0.326	0.322	0.318	0.314	0.310	0.306	0.303	0.299	0.295	0.291
1.0	0.288	0.284	0.280	0.277	0.273	0.269	0.266	0.262	0.259	0.255
1.1	0.252	0.249	0.245	0.242	0.239	0.235	0.232	0.229	0.226	0.223
1.2	0.219	0.216	0.213	0.210	0.207	0.204	0.201	0.198	0.195	0.193
1.3	0.190	0.187	0.184	0.181	0.179	0.176	0.173	0.171	0.168	0.165
1.4	0.163	0.160	0.158	0.155	0.153	0.150	0.148	0.146	0.143	0.141
1.5	0.139	0.137	0.134	0.132	0.130	0.128	0.126	0.124	0.121	0.119
1.6	0.117	0.115	0.113	0.111	0.110	0.108	0.106	0.104	0.102	0.100
1.7	0.098	0.097	0.095	0.093	0.092	0.090	0.088	0.087	0.085	0.083
1.8	0.082	0.080	0.079	0.077	0.076	0.074	0.073	0.072	0.070	0.069
1.9	0.068	0.066	0.065	0.064	0.062	0.061	0.060	0.059	0.058	0.056
2.0	0.055	0.054	0.053	0.052	0.051	0.050	0.049	0.048	0.047	0.046
2.1	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
2.2	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.030	0.029
2.3	0.029	0.028	0.027	0.027	0.026	0.025	0.025	0.024	0.024	0.023
2.4	0.023	0.022	0.022	0.021	0.020	0.020	0.019	0.019	0.019	0.018
2.5	0.018	0.017	0.017	0.016	0.016	0.016	0.015	0.015	0.014	0.014
2.6	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011	0.011
2.7	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008	0.008
2.8	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006	0.006
2.9	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005
3.0	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.003
3.1	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.002
3.2	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
3.3	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001
3.4	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.5	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.6	0.001	0.001	0.001	0.001	0.001	0.001	0.000			

These results show that the roots are bracketed, and good estimates of μ and σ may be obtained by interpolation between the sums. Interpolating between 1.2 and 1.3 using 2.09, 0, -5.20, we find $m_3=1.23$, and similarly $s_3=.185$. By doing two more calculations similar to the two illustrated above, one would verify the third figures in m and s and obtain good estimates for the fourth figures. Here, the results to three figures are $m=1.21$ and $s=.187$. However, the data do not

TABLE II

VALUES OF z/q

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.798	0.804	0.811	0.817	0.824	0.830	0.836	0.843	0.849	0.856
0.1	0.863	0.869	0.876	0.882	0.889	0.896	0.902	0.909	0.916	0.923
0.2	0.929	0.936	0.943	0.950	0.957	0.964	0.970	0.977	0.984	0.991
0.3	0.998	1.005	1.012	1.019	1.026	1.033	1.040	1.047	1.054	1.062
0.4	1.069	1.076	1.083	1.090	1.097	1.105	1.112	1.119	1.126	1.134
0.5	1.141	1.148	1.156	1.163	1.171	1.178	1.185	1.193	1.200	1.207
0.6	1.215	1.222	1.230	1 237	1.245	1.253	1.260	1.268	1.275	1.283
0.7	1.290	1.298	1.306	1.313	1.321	1.329	1.336	1.344	1.352	1.360
0.8	1.367	1.375	1.383	1.391	1.399	1.406	1.414	1.422	1.430	1.438
0.9	1.446	1.454	1.461	1.469	1.477	1.485	1.493	1.501	1.509	1.517
1.0	1.525	1.533	1.541	1.549	1.557	1.565	1.573	1.581	1.590	1.598
1.1	1.606	1.614	1.622	1.630	1.638	1.646	1.655	1.663	1.671	1.679
1.2	1.687	1.696	1.704	1.712	1.720	1.729	1.737	1.745	1.754	1.762
1.3	1.770	1.779	1.787	1.795	1.804	1.812	1.820	1.829	1.838	1.846
1.4	1.854	1.862	1.871	1.879	1.888	1.896	1.905	1.913	1.922	1.930
1.5	1.938	1.947	1.955	1.964	1.972	1.981	1.990	1.998	2.007	2.015
1.6	2.024	2.033	2.041	2.050	2.058	2.067	2.076	2.084	2.093	2.102
1.7	2.110	2.119	2.128	2.136	2.145	2.154	2.162	2.171	2.180	2.188
1.8	2.197	2.206	2.215	2 223	2.232	2.241	2.250	2.258	2.267	2.276
1.9	2.285	2.294	2.303	2.311	2.320	2.329	2.338	2.346	2.355	2.364
2.0	2.373	2.381	2.390	2.399	2.408	2.417	2.426	2.435	2.444	2.453
2.1	2.462	2.470	2.479	2.488	2 497	2.506	2.515	2.524	2.533	2.542
2.2	2.551	2.560	2.569	2.578	2.587	2.596	2.605	2.614	2.623	2.632
2.3	2.641	2.650	2.659	2.668	2.677	2.687	2.696	2.705	2.714	2.723
2.4	2.732	2.741	2.750	2.759	2.768	2.777	2.786	2.795	2.805	2.814
2.5	2.823	2.832	2.841	2.850	2.859	2.868	2.878	2.887	2 896	2.905
2.6	2.914	2.923	2.932	2.942	2.951	2.960	2.969	2.978	2.987	2.997
2.7	3.006	3.015	3.024	3.033	3.043	3.052	3.061	3.070	3.079	3.089
2.8	3.098	3.107	3.116	3.126	3.135	3.144	3.153	3.163	3.172	3.181
2.9	3.190	3.200	3.209	3.218	3.227	3.237	3.246	3.255	3.265	3.274
3.0	3.283	3.292	3.302	3.311	3.320	3.330	3.339	3.348	3.358	3.367
3.1	3.376	3.386	3.395	3.404	3.413	3.423	3.432	3.441	3.451	3.460
3.2	3.470	3.479	3.488	3.498	3.507	3.516	3.526	3.535	3.544	3.554
3.3	3.563	3.573	3.582	3.591	3.601	3.610	3.620	3.629	3.638	3.648
3.4	3.657	3.667	3.676	3.685	3.695	3.704	3.714	3.723	3.732	3.742
3.5	3.751	3.761	3.770	3.780	3.789	3.799	3.808	3.817	3.827	3.836
3.6	3.846	3.855	3.865	3.874	3.884	3.893	3.902	3.912	3.992	3.931
3.7	3.940	3.950	3.959	3.969	3.978	3.988	3.997	4.007	4.016	4.026
3.8	4.035	4.045	4.054	4.064	4.073	4.083	4.092	4.102	4.111	4.121
3.9	4.130	4.140	4.149	4.159	4.169	4.178	4.188	4.197	4.206	4.216
4.0	4.226	4.235	4.245	4.254	4.264	4.273	4.283	4.292	4.302	4.312

warrant any more accuracy in the roots than is given by m_3 and s_3, and one would not do the two extra computations. The results in Figure 5 were obtained by using the same set of observations (with mean 1.312 and standard deviation .158) as was used to obtain the results of Figure 1.

REFERENCES

[1] C. I. Bliss, *Annals of Applied Biology,* XXII, 1935, p. 134.

[2] R. A. Fisher, *ibid.*, p. 149.

[3] J. O. Irwin, *Journal of the Royal Statistical Society* (Supplement), IV, 1937, p. 1.

[4] J. O. Irwin and E. A. Cheeseman, *ibid.*, VI. 1939, p. 174.

[5] R. A. Fisher and F. Yates, *Statistical Tables*, 1938, Oliver and Boyd, London.

[6] F. Garwood, *Biometrika,* XXIII, 1941, p. 46.

[7] M. S. Bartlett, *Biometrics* III, 1947, p. 39.

[8] M. S. Bartlett, *Journal of the Royal Statistical Society* (Supplement), VIII, 1946, p. 113.

[9] P. C. Mahalanobis, S. S. Bose, P. R. Roy, and S. K. Banerji, *Sankhyā* I, 1934, p. 289.

A SHORT-CUT METHOD OF FITTING A LOGISTIC CURVE

WILLIAM A. SPURR AND DAVID R. ARNOLD
Stanford University

"Growth" curves have been used in many fields despite the extensive calculations required. These curves may be applied even more widely if their construction can be simplified. This article shows how to fit a logistic (Pearl-Reed) curve quite simply by means of a nomograph (which determines the upper limit from three selected points) and a logistic grid (which reduces the curve itself to a straight line).[1] Other "growth" curves may be fitted by constructing appropriate graphs.

THE METHOD

THE FIRST STEP in fitting a logistic curve to a time series by any method is to determine whether the process represented has the characteristics of population growth that will justify the use of the Pearl-Reed equation as a logical approximation.[2] The next step is to plot the complete series on semi-logarithmic graph paper to determine empirically whether the rate of growth is steadily declining. Then three points, sometimes geometric averages of several years, may be selected, representing typical levels of the early, middle, and recent periods. The points must be equidistant in time.

The nomograph (Figure 1) is then used to determine the upper limit k of the logistic curve from three selected points Y_0, Y_1, and Y_2 as follows: Compute Y_1/Y_0 and Y_2/Y_1 by slide rule;[3] place a transparent ruler across these values on the left and center scales, and read off k/Y_0 from the right hand scale. Multiply by Y_0 to find k.

Once the upper limit k is found, the Pearl-Reed curve may be determined readily from the logistic grid (Figure 2) which reduces this curve to a straight line. Divide the three selected points by k and plot these as percentages on the grid, using any convenient time scale on the X axis. Draw a straight line through the points[4] and read the values of this line at intervals of time, including the future, if a forecast is desired. Multiply these values by k, plot on the original semi-logarithmic chart and connect by a smooth curve. This is the logistic curve.

---

[1] The writers are indebted to Professors Albert Bowker and Alfred Niles of Stanford University for suggestions on this paper.

[2] See R. Pearl and L. J Reed, *The Biology of Population Growth*, Knopf, 1925.

[3] Y_2/Y_1 must be less than Y_1/Y_0 for this type of curve.

[4] If the three points do not fall on a straight line, an arithmetic error has been made.

These graphs may be used in any of four ways:

(1) to obviate all mathematical calculations except for several slide-rule ratios;

(2) to test any three selected points for goodness of fit. If the graphic logistic curve fits the data poorly, other combinations of three points may be tried. Then k may be computed mathematically for the points finally chosen;

(3) to serve as a quick check on the accuracy of machine calculation, since any error is likely both to distort k considerably and to destroy the linearity of the three points on the logistic graph;

(4) to make a further test of the validity of the logistic equation itself if desired. The raw data multiplied by $1/k$ must be roughly linear when plotted on the logistic grid to justify the use of this equation.

The following procedure is recommended: Use the nomograph for purposes (2) and (3) above, but also compute $1/k$ mathematically, since the nomograph is fairly approximate. Then fit the actual curve by means of the logistic grid alone, since this is probably quite accurate enough. Enlarge Figures 1 and 2 by photostat to a convenient size, and use tracing paper over the latter.

DISCUSSION

The logistic, Gompertz, and other S-shaped curves representing retardation of rate of growth have come into wide use since 1920 as means of measuring and projecting biological growth generally, and secular trends of population and industries in particular. The simple logistic or Pearl-Reed curve fitted to three selected points is probably the most widely used of these curves,[5] so it has been selected for solution here.

The logistic curve may be written:

$$Y = \frac{k}{1 + e^{a+bX}}$$

where the curve ordinate Y approaches k the upper limit as time, X increases; e is the natural logarithm base 2.71828, and a and b are constants, b being negative.

Even this simple form of the growth curve, however, requires the lengthy calculation of the three constants, k, a, and b, a procedure that must be repeated for each set of three points that are tested for goodness of fit[6]. Moreover, despite the precision of mathematical calculation, the results are only approximate, since (1) the equation itself is merely one of several empirical approaches to the "law of growth", (2) the selection of three points is subjective, and (3) errors of any kind are exaggerated in extrapolating to find the upper limit.

---

[5] See Simon Kuznets, *Secular Movements in Production and Prices*, pp. 59–68, 197–199, Houghton Mifflin, 1930, or a recent popular description in E. R. Dewey and E. F. Dakin, *Cycles*, Holt, 1947, Ch. I–IV.

The graphic method obviates the following calculations:

$$k \quad \begin{matrix} 2Y_0Y_1Y_2 - Y_1^2(Y_0 + Y_2) \\ Y_0Y_2 - Y_1^2 \end{matrix}$$

$$= \log_e \frac{k - Y_0}{Y_0}$$

$$\frac{1}{n}\log_e\left(\frac{Y_0(k - Y_1)}{Y_1(k - Y_0)}\right)$$

$$Y = \frac{k}{1 + e^{a+bX}}$$ (This must be calculated for each point of the fitted curve.)

The errors of graphic measurement are believed to be small compared with the errors implicit in the Pearl-Reed equation itself and the selected points as described above.

THE NOMOGRAPH

The nomograph for finding the upper limit was constructed by reducing the logistic equation to:

$$\theta B - \theta A - 2AB + A^2B + A = 0$$

where $\theta = k/Y_0$, $A = Y_1/Y_0$ and $B = Y_2/Y_1$.

This equation was expressed as a third-order determinant which was multiplied by a matrix of transformation to produce a nomographic determinant.[7]

The equations for the coordinates of the nomographic scales (in arbitrary linear units) were then found to be:

---

[6] See F. E. Croxton and D. J. Cowden, *Applied General Statistics*, pp. 452–458, Prentice-Hall, 1939. The methods of least squares and moments are much more cumbersome than that described

The present method appears to be shorter than that recently suggested by Dudley J. Cowden, (*Journal of the American Statistical Association*, Dec. 1947, pp. 585–590), since: (1) the trial values of k may be determined from a nomograph, rather than computed; (2) the logistic grid straightens out the logistic curve itself, whereas Professor Cowden uses a semi-logarithmic grid to produce a linear transformation of the function $\frac{1}{Y_c} - \frac{1}{K}$, which requires five extra columns of computations or scale readings (p. 589). The method described here should also be shorter than that of Raymond Pearl (Introduction to Medical Biometry and Statistics, Third Edition, Chap. XVIII, W. B. Saunders Co., Philadelphia, 1940) for similar reasons.

[7] See F. T. Mavis, *The Construction of Nomographic Charts*, pp. 56–57, Scranton, Pennsylvania: International Textbook Co., 1939.

$$X_\theta = 0 \qquad Y_\theta = \frac{\theta - 10}{4}$$

$$X_B = \frac{2.52B}{0.76B - 4} \qquad Y_B = \frac{5.76B - 9}{4 - 0.76B}$$

$$X_A = \frac{2.52A}{0.76A - 4} \qquad Y_A = \frac{7.76A - A^2 - 10}{4 - 0.76A}$$

A nomograph of any size may be constructed from these equations by using an appropriate scale. However, it is simpler merely to enlarge Figure 1 by photostat or photograph, extending the k/Y_0 line upward if desired.

The error in reading k from the nomograph is on the order of 2 per cent for the usual range of k values shown in Figure 1 ($4 < k/Y_0 < 45$) using a 7″ by 10″ chart, but is reduced for all points on the logistic curve since they are only fractions of k. These errors are fairly small compared with the hyper-sensitivity of k itself to the three subjectively chosen points.[8] The error increases for the extremely high and low values of k which are not shown. These should be computed mathematically.

The nomograph in general should be more generally used in statistics, as it is in engineering. It has these advantages: (1) the chart is simple and easy to read; (2) interpolation is made fairly accurately along a scale; and (3) the chart shows at any point the nature of the change in one variable due to changes in the other variables.[9] A curved nomograph, such as this one, saves even more labor than the more usual linear type,[10] since it replaces a more complex equation with the attendant labor of calculation.

In case it is desired to fit a logistic curve by using all available points rather than only three points, the following short-cut method is suggested for approximating the upper limit k. This method utilizes the principle that the per cent rate of increase in a logistic curve is a linear

---

[8] For example, if the first two points are 1 and 2, a variation in the third point from 3 to 4 will cause k to vary from 4 to infinity. Estimates of the upper limit of California's population (which is nearly an exponential curve) by 18 Stanford students based on various selections of three points from the same data varied from 18 to 613 millions. The variation is relatively small, though still considerable, in an older series. According to M. C. Rorty (*Journal of the American Statistical Association*, March, 1931, pp. 8–9) the growth curve is significant when the series has reached within 20 to 30 per cent of the saturation point.

[9] Joseph Lipka, *Graphical and Mechanical Computation*, p. 44 f, Wiley, 1918.

[10] See J. W. Dunlap and A. K. Kurtz, *Handbook of Statistical Nomographs, Tables, and Formulas*, Part I, World Book Co., Yonkers, New York, 1932, for an excellent compilation of linear statistical nomographs.

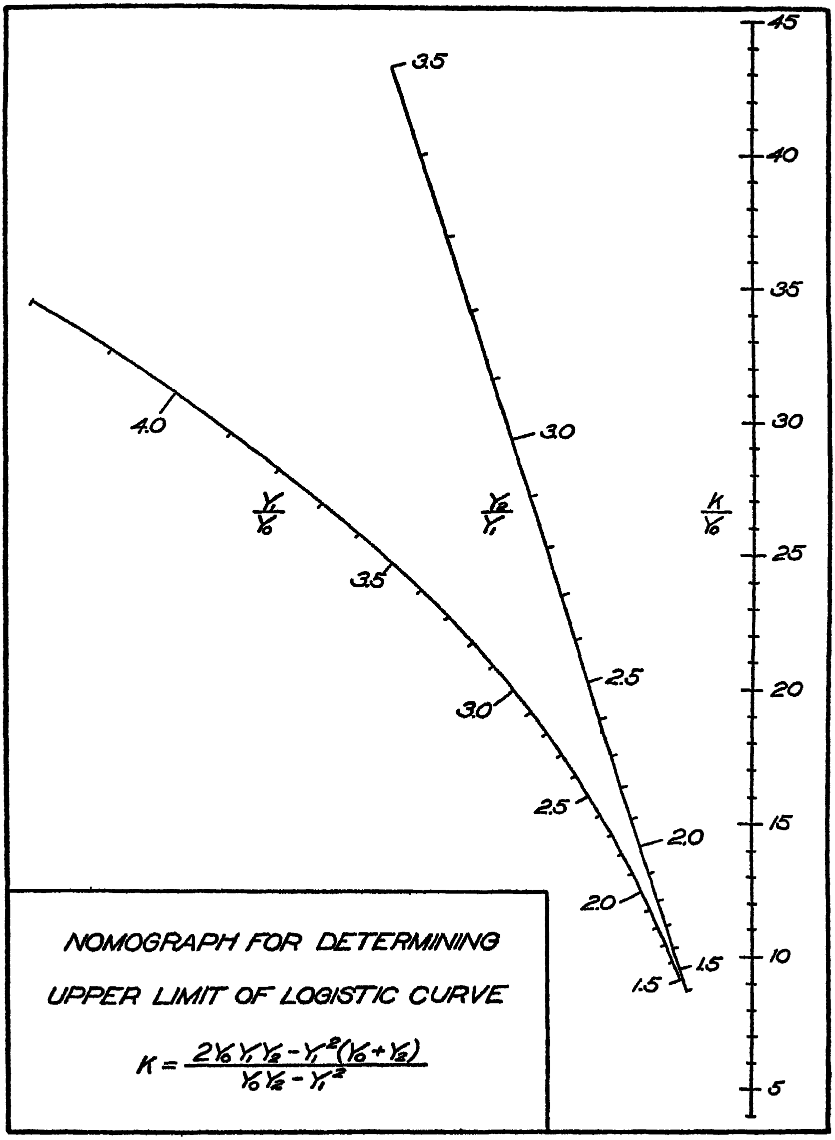

3.5
3.0
2.5
2.0
1.5
$\frac{Y_2}{Y_1}$
4.0
3.5
3.0
2.5
2.0
1.5
$\frac{Y_1}{Y_0}$
45
40
35
30
25
20
15
10
5
$\frac{K}{Y_0}$
NOMOGRAPH FOR DETERMINING
UPPER LIMIT OF LOGISTIC CURVE
$K = \frac{2Y_0Y_1Y_2 - Y_1^2(Y_0+Y_2)}{Y_0Y_2 - Y_1^2}$

FIGURE 1

function of the population itself, and that the rate approaches zero as the population approaches k. Compute by slide rule the percentage increases in the data from point to point. Then estimate the geometric mean of each pair of points from the middle of the line connecting the points on semi-logarithmic paper.[11] Now plot the percentage increases on the Y axis against the corresponding means on the X axis in a scatter diagram. Fit a straight line to these points by inspection or by least squares. (The goodness of fit of this line is a test of the validity of the logistic.) This line extended will cross the X axis at the desired value of k. However, as Yule said,[12] "I do not think this is a very good method of fitting." The graphic form of the method described above is shorter than Yule's or Hotelling's versions and is about as accurate as Yule's; but it still involves several approximations, and is much more work than the use of the nomograph with three points only.

THE LOGISTIC GRID

The logistic grid not only provides a linear transformation of the logistic curve, but may also be used to find the remaining constants in its equation, if desired, as follows:

$$b = -\frac{1}{X_{50\%} - X_{26.89\%}}$$

$$a = -bX_{50\%}$$

where the X values are the abscissas corresponding to the Y/k values of 50 per cent and 26.89 per cent respectively.[13]

Any logistic scale may be computed as follows: For all desired scale values of Y/k up to 50%, list the corresponding values of $\log_e(k/Y-1)$

---

[11] More precise but laborious methods of computing these values have been suggested by Yule (*Journal of the Royal Statistical Society*, 1925, pp. 1–58) and Hotelling (This *Journal*, September, 1927, pp. 283–314). Hotelling shows an illustrative scatter diagram of this general method on page 299.

[12] *Op. cit.*, Method 3, p. 52.

[13] Proof: 1. Let $Y/k=50\%$ in the logistic equation. Then $a=-bX_{50\%}$

2. Let $Y/k=p$. Then $a+bX=\log_e\left(\frac{1-p}{p}\right)$. Substituting a above,

$$b=\frac{\log\left(\frac{1-p}{p}\right)}{Xp-X_{50\%}}.\ \text{Let}\ \log\left(\frac{1-p}{p}\right)=1.\ \text{Then}\ p=\frac{1}{1+e}=26.89\%$$

$$\therefore b=\frac{1}{X_{50\%}-X_{26.89\%}}$$

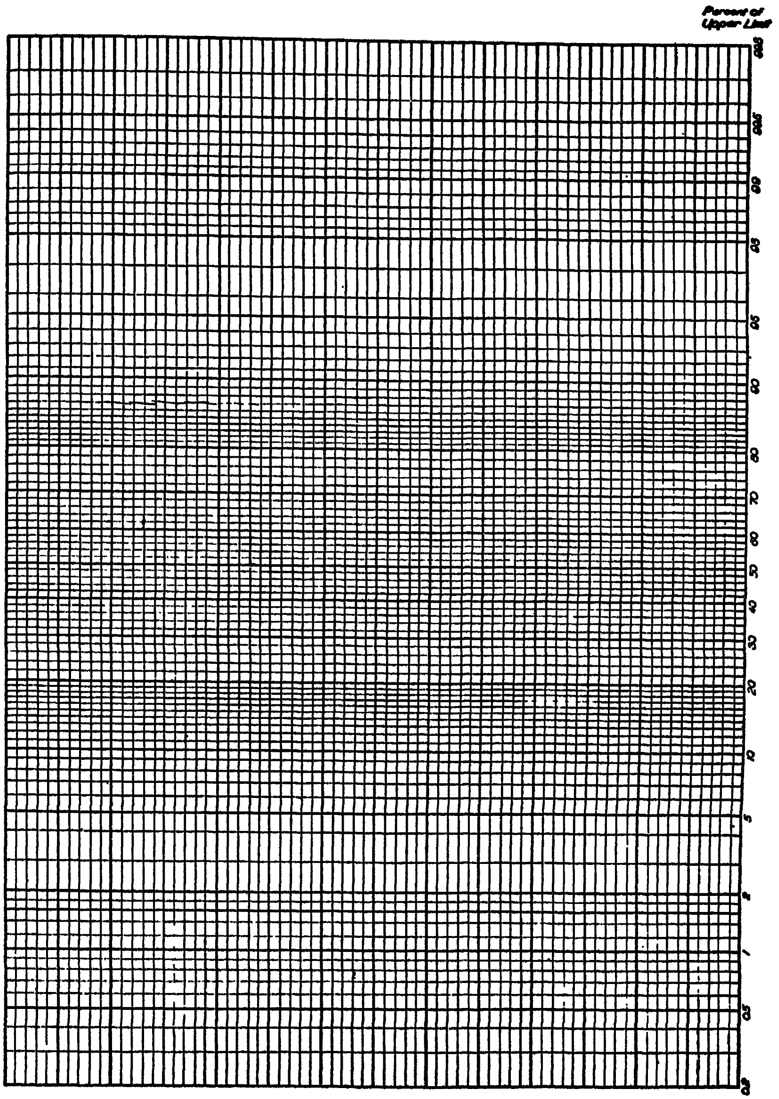


FIGURE 2

from tables of reciprocals and natural logarithms (such as *Barlow's Tables* and the Federal Works Agency, *Tables of Natural Logarithms*). Multiply the results by a constant to obtain the desired scale. Then measure these figures as linear distances down from the center of the chart for Y/k values below 50 per cent and up for complementary values above 50 per cent. For example, for $Y/k = .01$ per cent (or 99.99 per cent) $k/Y - 1 = 9999$, its $\log_e$ is 9.210, so this line would be scaled 9.210 inches or a multiple thereof below (or above) the middle of the chart.[14]

Again, however, a photostatic enlargement of Figure 2 is a simpler method of constructing this scale.

The accuracy of reading the logistic graph varies with the scale and with the vertical position of points on the chart. For a 7″ by 10″ chart, the average error will vary from about .01 of 1 per cent at the top and bottom to about 0.5 of 1 per cent at the center. The error will be proportionally less for larger charts. Again, these errors seem small compared with the approximations involved in the selection of three points and the equation itself. The logistic scale, unlike the nomograph, is valid for all values of k.

The logistic grid superficially resembles the probability grid which provides a linear transformation of the cumulative normal probability curve, derived from the law of errors rather than the "law of growth." The use of the logistic scale will avoid the sizeable error involved in the use of the probability scale as a means of approximating the Pearl-Reed curve.[15]

[14] Linear ordinates (τ) for integral percentages of Y/k (y') from 1 per cent to 99 per cent are listed by Yule, *op. cit.*, p. 48.

[15] For example, Croxton and Cowden, *op. cit.*, pp. 458–461, reach an upper limit of 400 millions for U. S. population from probability paper, compared with 190.8 millions from the Pearl-Reed equation and their consensus of "well under 200 millions" from other sources. The logistic function is claimed to be superior to the probability function as a measure of growth by C. P. Winsor (*Journal of the Washington Academy of Sciences*, February 19, 1932, pp. 73–84) and by J. Berkson (*Journal of the American Statistical Association*, September, 1944, p. 357 f). It is criticized by A. L. Bowley (*Journal of the Royal Statistical Society*, 1925, pp. 76–81), G. R. Davies (*Journal of the American Statistical Association*, 1927, pp. 370–374), and G. H. Knibbs (*Journal of the American Statistical Association*, 1926, p. 381 f and 1927, p. 49 f).

BOOK REVIEWS

Edited by
OSCAR KRISEN BUROS
Rutgers University

Bevezetés a Statisztika Tudományába: Rész I. (Introduction to the Science of Statistics, Part 1.) *Gyula Abay.* Pécs, Hungary: Rákóczi Nyomda (Rákóczi-Út 35), 1945. Pp. ii, 98.

Statisztika Módszertani Alapvetés. (Elements of Statistical Methodology.) *Loránd Schweng.* Budapest, Hungary: Stephaneum Nyomda, 1944. Pp. xviii, 366.

REVIEW BY EUGENE LUKACS
Professor of Mathematics, Our Lady of Cincinnati College
Cincinnati, Ohio

ABAY'S textbook, intended for students of law and social sciences, is divided into five chapters. Chapter 1 surveys the various fields of applied statistics and classifies social statistics into statistics involving population problems, economic statistics, and statistics related to public health and education. Chapter 2 discusses the history of statistics with particular reference to Hungarian statistics. Chapter 3 treats the gathering of statistical data and the work of the official Hungarian statistical services. Chapter 4 discusses the organization of the data and the construction of empirical frequency distributions; numerous examples and graphs are given. Chapter 5 deals with measures of central tendency and dispersion.

Although Abay states that it is impossible to discuss modern statistics without using mathematical methods, he avoids the use of mathematical formulae. There are a few incorrect statements: for example, on pages 9 and 10, the author states that every frequency curve would be a normal curve if one only had enough data. Statistical methods developed within the past forty years are not considered. The book is not likely to lead the student to a proper appreciation of statistical methods.

Schweng's text—prepared for students of social sciences, especially economics—discusses in great detail methods of collecting and organizing statistical data. The first part of the book deals with the collection of data and the nature of statistical tables. The proper consideration of significant figures is greatly emphasized. The second part discusses graphs, diagrams, ratios, and percentages. The third part presents an elementary discussion of frequency distributions and measures of central tendency and dispersion. Very few formulae are used, and no proofs are given. The normal distribution is barely mentioned. Parts 4 and 5 give brief introductions to time series and index numbers. Part 6 surveys the most important Hungarian index numbers.

Numerous examples are given throughout the text. Statements and definitions are always illustrated by tables and graphs. The 5-page bibliography lists a large number of American, British, French, German, and Italian books. The book does not reflect the progress of theoretical statistics during the last forty years and therefore can not be considered an introduction to modern statistical methods. Nevertheless, the book may prove useful for the consumer of Hungarian statistics rather than for the producer of statistical data.

College Algebra. *A. Adrian Albert* (Professor of Mathematics, University of Chicago). New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1946. Pp. xii, 278. $2.75. (London W.C. 2: McGraw-Hill Publishing Co. Ltd. [Aldwych House, Aldwych]. 14s.)

REVIEW BY PAUL S. DWYER
Professor of Mathematics
University of Michigan, Ann Arbor, Michigan

THIS book incorporates a drastic revision of the conventional approach and treatment of the subject matter of college algebra and as such deserves the attention of statisticians. Important objectives of the author are:

a) The material should be presented as a unified and compact body of mathematical theory. As worked out by the author, this unity is achieved by the successive study of such general topics as the number systems of elementary mathematics, polynomials and allied functions, algebraic identities, equations and systems of equations. Considerable material not usually presented in a course in college algebra is included in this approach.

b) The definitions and theorems should be stated accurately. This means, of course, that the author has been quite formal in the presentation of the material. This formality will probably be approved by mathematical statisticians. It may not be so welcome to other statisticians who wish to study the main results with as little of the vocabulary and notation of the mathematician as is possible.

c) Proofs of results should be given only when it is probable that the better students will understand them and when their inclusion will add to the understanding of the results. The author is very careful to indicate each statement which is lacking in proof. In this connection, the final chapter (Chap. 10) on matrices and quadratic forms should be noted since the author makes no pretense at proofs here. This chapter, frankly experimental, is for the purpose of providing matrix material for those who understand the definitions, notation and statements of theorems. As such it will be of interest to many statisticians.

d) There should be an adequate number of illustrative examples and exercises. The author has provided these and it seems that he has provided another feature which will be of interest to applied statisticians—a greater emphasis upon numerical work than is found in most texts on college algebra.

e) The text should abound with additional material on the same subject for the better student. The author has accomplished this with the use of additional sections and chapters which are indicated as "full course" and which tend to amplify the minimum treatment. In some of this optional material the author attempts to relate the subject matter of college algebra to other material in college mathematics. Chapter 8, "Vectors in the Plane" is of this sort.

The statistician will be particularly interested in the treatment of Chapter 9, "Matrices, Determinants and Linear Systems," and Chapter 10, referred to above. He may or may not approve of the introduction of the concept of a matrix before that of a determinant, but he certainly will approve of the use of elimination methods, rather than determinants, in solving linear simultaneous equations. It is unfortunate that the author does not also show that elimination (condensation) methods can be used in evaluating determinants and, in fact, that the same general procedures can be employed in the forward solution no matter whether one solves equations, evaluates determinants, or computes the inverse of a matrix. The author does give a general treatment of determinants of any order.

Chapter 10 has been described above. It gives a number of basic theorems and techniques which are commonly useful, together with simple problems. It is unfortunate that the definition of matrix multiplication is given only in terms of the formal row-by-column multiplication of A and B and is not extended to the frequently useful column-by-column multiplication of A' and B or the row-by-row multiplication of A and B'.

The first copies of the book contained an unusually large number of mistakes not only in the answers to the problems but in the theory itself. Most of these have been removed in the second printing.

The author is to be commended for the way in which he has accomplished his objectives. Each reader must determine for himself whether these are appropriate objectives for him. I think that we as statisticians will agree that the emphasis upon numerical work, the more practical approach to the problem of simultaneous linear equations, and the inclusion of matrix material are all steps in the right direction.

A Concise Manual of Statistics: With Special Reference to the Requirements of Students for Municipal Examination, Second Edition. *Clement Burton* (Borough Treasurer's Department, Hackney Town Hall, London E. 8, England). London E.C. 2: Gee and Company (Publishers) Limited (27–28 Basinghall St.), 1946. Pp. 182. 15s.

REVIEW BY HARRY PELLE HARTKEMEIER

Professor of Business Statistics, University of Missouri

Columbia, Missouri

THE preface of this book contains the statement: "The increasing importance of the subject of statistics in commercial and municipal activity has prompted the preparation of this handbook in the hope that it will prove of interest generally, and will be of assistance to those who are studying for the examinations of the various professional bodies of accountants, and more particularly for those of the Institute of Municipal Treasurers and Accountants (Incorporated)." In England it is necessary for a person to pass an examination in the field of statistics before he is allowed to practice with something equivalent to our C.P.A. This is an excellent practice and the authorities responsible are to be congratulated. It would be a very good idea if accountants in the United States were required to demonstrate a similar knowledge of statistics before they are recognized as Certified Public Accountants. Most accountants are deficient in a knowledge of sampling techniques and tests of significance or reliability.

This book is an excellent one in view of its purpose as a manual or handbook. It will serve an accountant well when he wants to refresh his memory about a statistical formula or process. Presumably, the accountant has already known something about statistics and needs primarily a rapid review. This book is not so well suited to be an elementary textbook for a new student to use to obtain an introduction to statistics. One book could hardly be both an elementary textbook and a concise manual or handbook.

The book seems a bit jerky at some points. One gets the impression that it was written a little longer and parts were deleted to shorten the book. This could have been the work of some editor not familiar with the material or subject. At most places the material reads smoothly and for an American reader the book has some delightfully surprising expressions. For example, rounded off data are items "shorn of digits."

Everyone has the right to select his own symbols, but the American reader is not apt to like the choice of *a* for the arithmetic average, Z for the mode, and M for the median. The usual meaning for the probable error is found on pages 123 and 124 in connection with the explanation of correlation, but there may be some confusion when the reader compares this meaning with the statement on page 29, "When figures are taken, for example, to the nearest 100, the maximum possible error is 49; but as the error may be anywhere between 0 and 49 the average or probable error is 24.5."

On page 124 the reader encounters eight statements dealing with the interpretation of the value of *r*. Two of them are: "If *r* is less than .3—little evidence of correlation." "If *r* is more than 3 times the probable error—evidence of correlation." No mention is made of the possibility that *r* could easily be 3 times the probable error *and* be less than .3. The two different conclusions conflict. What could be said about such a situation is that when *r* is more than 3 times the probable error there is reliable evidence that some correlation exists but when *r* is also less than .3 the correlation is not large enough to be of practical value for estimating the size of one variable from a knowledge of the other variable and the equation of relationship.

Very careful type setting or very good proofreading was responsible for the fact that very few errors are noticed.

One of the principal virtues of the book is the honest recognition of the value of the work of the statistician. "The position of the statistician lies between the accountant or cost accountant and the economist or executive authority." No doubt this will sound like heresy to accountants in the United States, who for years have claimed the position of being the right-hand man of the executive. At long last, people are beginning to realize that accounting is but a part of the much broader science of business statistics.

Mathematische Erblichkeitsanalyse von Populationen. *Gunnar Dahlberg* (Professor and Head of the Swedish State Institute for Human Genetics and Race Biology, Stockholm, Sweden). Aus Dem Staatlichen Institut fur Menschliche Erblichkeitsforschung und Rasenbiologie zu Uppsala. Diese Arbeit erscheint auch als Supplementum CXLVIII von Acta Medica Scandinavica. Uppsala, Sweden: Almqvist & Wiksells, 1943. Pp. 219. Paper.

REVIEW BY OYSTEIN ORE
Professor of Mathematics, Yale University
New Haven, Connecticut

IN THIS work the author gives a review of the applications of mathematical methods to certain problems of inheritance. Particular emphasis is placed on the consequences for human society of some of the mathematical theories and hypotheses and almost all illustrations are drawn from this sphere.A basic knowledge of genetical terminology is presupposed. The introductory sections are devoted to a discussion of the use of the elements of probability to the determination of the various gene combinations and their frequencies as they may occur in a population. In the later sections considerations and methods of an actuarial nature predominate. Here one finds an extensive analysis of the effect which certain conditions in the population, for instance occurrence of mutations, existence of isolated groups and selection of various kinds, e.g. mating with closely related individuals, will eventually have on the gene frequencies in a society. Of particular value are the numerous tabular computations which illustrate the consequences of the various numerical assumptions one may make in regard to the perturbing factors.

Cycles: The Science of Prediction. *Edward R. Dewey* (Director, Foundation for the Study of Cycles, 274 Madison Ave., New York, N. Y.) and *Edwin F. Dakin* (Counsel in Public Relations, Plymouth, N. H.). New York 10: Henry Holt & Co. (257 Fourth Ave.), 1947. Pp. xi, 255. $3.00. *Two reviews follow.*

REVIEW BY MILTON FRIEDMAN
Associate Professor of Economics, The University of Chicago
Chicago, Illinois

THERE are, say Dewey and Dakin, a number of important rhythms or cycles in economic activity. These cycles have fixed periods and are common to many different activities. Any particular economic series is to be conceived as the sum of such cycles superimposed on a trend and disturbed—though only temporarily—by such erratic and random factors as wars. The appropriate technique for predicting the future course of any economic series is to determine empirically from its past behavior its trend and the chief cycles it contains, to synthesize these, and extrapolate them into the future. Dewey and Dakin conclude from analysis of this type so far performed that the four most important of the many rhythms have been isolated in our economy are:

> A. The 54-year rhythm in wholesale prices and industrial innovations.
> B. The 9-year rhythm in wholesale prices, security prices, pig iron production, and industrial activity.
> C. The 3½-year rhythm in security prices and in business activity, wholesale and retail.
> D. The 18-year rhythm in real estate activity, and in related industrial enterprise (p. 135).

These rhythms operate in an economy that is growing at a declining rate and is approaching "maturity," so that their downward phases will appear relatively more severe, and their upward phases less buoyant, in the future than in the past. The 54-year rhythm turned down in 1925, and "is due to reach bottom in 1952" (p. 226); the 9-year rhythm was due for a peak in 1946 and a subsequent decline until 1951; the 3½-year rhythm was due for a peak in 1947 and again in 1950 and for lows in 1948 and 1951; the 18-year rhythm in real estate activity "apparently reached its high about late 1942" and is "due to decline to a low around 1953" (p. 227). "The foregoing is not meant to suggest that our rhythms indicate collapse of our economy. We may only infer, if they continue, that around 1947 a fall-off will start, and our economy will experience rather protracted declines in prices, production, employment, and economic activity generally, reaching a bottom that would presumably be dated sometime in the early fifties" (p. 230).

The book presenting this message is difficult to classify. It is not a scientific book: the evidence underlying the stated conclusions is not presented in full; data graphed are not identified so that someone else could reproduce them; the techniques employed are nowhere described in detail. Yet it is not a popular book: there are too many graphs and too much technical jargon to call it that. Its closest analogue is the modern high-power advertisement—here of book length and designed to sell an esoteric and supposedly scientific product. Like most modern advertising, the book seeks to sell its product by making exaggerated claims for it (that it will give sufficiently accurate forecasts to contribute substantially to success in business), showing it in association with other valued objects which really have nothing to do with it (established cycles in physical, meteorological, or other natural phenomena, and the successful use of cycles in predicting such phenomena as tides), keeping discreetly silent about its defects or mentioning them in only the vaguest form (no examples of past failures of cyclical patterns to reproduce themselves are mentioned), and citing authorities who think highly of the product. As with most modern advertising, the reader will find the book easy and interesting to read; when he finishes, he will find himself favorably inclined toward the product; yet, when he reflects, he will find that there is little reason why he should be. What is good is commonplace, and what is claimed as exceptional is unproven. There is little doubt that approximate rhythms exist in economic activities, that there is a building cycle in the United States about 18 years in length, and a business cycle about 3½ years in length. There is also little doubt—though the authors do not mention this—that every attempt to use these facts or any others to forecast economic activity has to date met with failure. Beyond this, we know little. The authors assert a good deal more, but these assertions, though illustrated by selective examples, are not demonstrated by any comprehensive or carefully presented body of evidence. Even the techniques used to derive the particular conclusions about lengths of cycles and the like are only alluded to in passing in a chapter (Chapter 11, "Analysis and Synthesis") which is clearly intended to impress the reader with the scientific magic used to get the stated results rather than to initiate him into the art or to enable him to test the results obtained by the authors.

An excellent, though somewhat extreme, example of the level of analysis in the book is its treatment of secular trends. The book makes much of the universality of retardation in the rate of growth of individual industries. From this well established phenomenon, it erroneously deduces that the economy at large must grow at a retarded rate and approach maturity. And it makes this erroneous deduction despite a reference (in which, incidentally, the title is given incorrectly) to A. F. Burns' *Production Trends in the United States Since 1870* as "required reading" (p. 52). Yet one of the most important conclusions of Burns' book is that retardation in the rate of growth of the economy as a whole is not a necessary result of retardation in the rate of growth of each industry separately; and, further, that there is little evidence in American data of the existence of retardation for the economy as a whole.

REVIEW BY MAX SASULY
Research Associate, The Robinson Foundation
Washington, D. C.

THIS is an important, certainly a significant book. We may discount the jacket statements that it

> establishes the ground for a completely new approach to the economic problems of our era. It puts a tool in the hands of businessmen, statesmen, and serious readers which will enable them to predict the economic outlook for the critical years ahead.

This is over-optimistic, even for a jacket blurb. But the book is significant as a contemporary sociological document, emphasizing the difficulties of developing a realistic economics in the service of society.

The authors, according to the jacket, are experienced in counseling "corporate enterprise." The book is in fact a model of skillful writing. A good deal of information is given about "cycles," but not more than confused intimations about the science of prediction. The content of the book boils down to the *quest* for a science of prediction—by business counselors. The title and the style are, however, very effective. Backed by skillful publicity in the preceding years, the book has been a best-seller for months! The objective of the reported research is not abstract theory; it is the age-old quest of "Forecasting for Profit." The book may contribute to this aim somewhat, indirectly. Indirectly it may contribute to the progress in cycle research by high-lighting the gulf between the demand for realistic economic analysis and the generally barren theory and disputation available as supply.

The book appears to be a progress report of "a non-profit corporation" organized as the Foundation for the Study of Cycles, of which Mr. Dewey is director. The early chapters treat of trends and patterns in typical industries. Well-known incidents of recent decades in commerce and industry are presented in the jargon of the financial and commodity marts with disarming easy simplicity. The next few chapters deal with rhythmic cycles. Certain cycles of nature, some more or less recently noticed, are first discussed. Then a chapter each is given to certain marked economic cycles that have become familiar in recent literature—the cycles of 54, 9, 3½, and 18 years. The final chapters treat of purported causal and correlation analyses of economic series; the effects of the War; and probable postwar trends and rhythms. In three short appendices are presented—as tools for the reader (?) —the rudiments of the "ratio" scale, moving averages, and the "section moving average." No index is supplied.

Very effective advance publicity was given the book in a popular article[1] in 1943. This led Smith and Duncan[2] to the belief—in 1944—that:

> So general was the interest in cyclical behavior [after the 1930's] that by 1940 the Foundation for the Study of Cycles . . . proposed to help in the task of integrating the work of the thousands of scientists and statisticians who are contributing in various fields to the study of cycles. Not only have cycles been found to exist in the realm of business activity, but scientists in many other fields believe they have discovered cyclical behavior in their respective studies.

Dewey stated, in his 1943 article, "the results are most encouraging. Forecasts made by projecting well-established cycles are coming true every day." He predicted:

> Figure on the possibility of a major financial and economic setback starting about 1947. . . . Heavy industry will begin to slack off next year [1944] if the 41-month rhythm of peaks and valleys prevails over this war as it did during the last one.

In the present book, in 1947, Dewey and Dakin purport to do more than

---

[1] Dewey, Edward R. "Science Predicts the Future." *Am Mag* 136:90–2 S '43.

[2] Smith, James G., and Duncan, Acheson J. *Elementary Statistics and Applications: Vol. I, Fundamentals of the Theory of Statistics*, p. 652. New York: McGraw-Hill Book Co., Inc., 1944.

"integrate" the work of numerous others. They believe that [like nuclear physics]

> economics is now reaching a point where it can hope also to make rather accurate predictions, within limits which this study will explain (p. vii).
>
> The reader will be introduced to a method of thinking about the future which—new though it may be to him—seems definitely to have proved of value. It is this method which is of fundamental importance—an importance greater than any specific conclusions to which it may lead. For on its validity depends the whole value of the conclusions (p. viii).

The authors then record their deep indebtedness

> to those whose names, equations and graphs line the pages of this book—and to many others unnamed—. . . . Theirs is the pioneering that is moving economics out of the blind alley where it stood for many years, so that it can take its rank as a true science (p. viii).

The substance of the book is in fact largely derived from available familiar studies on cycles: Persons, Mitchell, Schumpeter, and especially W. I. King on economic cycles; Harlan Stetson, and others on sunspot and other cycles in nature. The authors, however, read much more of an inexorable periodicity into the work of these well-known scholars than is warranted. The disturbing implication of this determinism is explained away (p. 98) by reasoning that so far has proved inscrutable to the reviewer. Indeed, elsewhere the authors aver blithely that social forces—at least quasi-volitional—can thwart, or reinforce the normal pattern of the unfolding cycle (p. 113)—this with reference to comments on the "inflationary forces" influencing us at present (1947). Some lingering doubts as to the consistency of their position are disposed of by a singular argument: one must not "reason solely on the ordinary cause-and-effect basis" in business, or in general economic analyses, (p. 114). The authors purport to derive support for this dialectic handspring from some of Eddington's lay reflections (in his Gifford lectures, 1925), and from nebulous references to Ouspensky, Spengler and Jeans.

The principal defect of the authors' discussion is a misreading of their primary sources, and misinterpretation of others. They also omit reference to outstanding works that quite invalidate their position. Thus, Harlan Stetson, a veteran physicist astronomer, discusses sunspots and their influence in a number of studies extending over twenty years, and clearly shows[3] that the 11-year cycle, so striking in appearance on the graph, varies widely in length—10 to 16 years—and has no more scientific validity than the periods of 23, 37, 77, 83, and more years found by other competent investigators. Mitchell and his coworkers have studied hundreds of economic series during the past three decades. In the last annual report[4] of the National Bureau of Economic Research, Burns demonstrates again that recurrences may always be expected, but, so far, they cannot be predicted. The repeated failures during the past century of hypotheses suggesting periods from forty months to sixty years in economic series are stressed anew by Burns (pp. 3–5). These cannot be blandly explained away by the Dewey-Dakin Chapter 13, on "Avoiding Some Economic Illusions." It is impressive, but neither relevant nor conclusive, when they quote in their argument this sententious aphorism from Spengler:

---

[3] Stetson, Harlan. *Sunspots and Their Effects*, pp. 162–72. New York: McGraw-Hill Book Co., 1930.

[4] Burns, Arthur F. *Stepping Stones Toward the Future*. Twenty-Seventh Annual Report of the National Bureau of Economic Research. New York: the Bureau, March, 1947.

> The separate sciences—epistomology, physics, chemistry, mathematics, astronomy—are approaching one another with acceleration, converging toward a complete identity of results.

Their case would not be bettered if, as the authors state, "he could well have added biology, psychology, economics, and sociology to his list" (p. 198).

From the technical statistical analysis side, the thesis of the authors is thoroughly repudiated in the extended recent study (not noted by them) by M. G. Kendall.[5] Kendall's authoritative analysis is supported by the panel of eminent English statisticians discussing his paper.

The discussion in the book ignores the well-considered, authoritative articles in the *Encyclopedia of the Social Sciences* ("Business Cycles" by Wesley C. Mitchell and "Forecasting Business" by Garfield V. Cox) and in the *Encyclopedia Britannica,* Fourteenth Edition ("Trade Cycle" by D. H. Robertson and "Trade Forecasts" by W. I. King). These authors rejected without qualification the idea that economic cycles can be predicted. Robertson, noting that for England crises occurred in 1835, 1847, 1866, 1873, 1882, 1890, 1900, 1907, 1913 and 1920, stressed that it is unsafe to dogmatize about the length or degree of phase synchronism in the leading industries or in leading industrial countries—for any cycle phase. W. I. King, on whose *The Causes of Economic Fluctuations* (1938) Dewey and Dakin lean heavily, noted clearly with regard to patterns that frequently repeat themselves, that the difficulty is "that in many cases expected repetition fails to materialize." The authors also make extensive use of Edgar Lawrence Smith's study, *Tides in the Affairs of Men* (1939). They give no inkling that he subtitled his book as merely "An Approach to the Appraisal of Economic Change." Finally, it is singular that the authors make no reference to Burns and Mitchell's significant book, *Measuring Business Cycles,*[6] which was available early in 1946. Preliminary versions of its content have been available for over ten years, e.g., Arthur F. Burns, in a 1936 paper,[7] reported that Mitchell's "forthcoming second volume on business cycles" exhibits the markedly wide range of variation in length of various economic cycles.

The wide variations in cycle length is the nub of the difficulty in the quest for means of *predicting* cycle phases. This difficulty has completely escaped the authors' notice. They draw heavily on Schumpeter's notable *Business*

---

[5] Kendall, M. G. "On the Analysis of Oscillatory Time-Series." *J Royal Stat Soc* 108:93–141 '45.

[6] Burns, Arthur F, and Mitchell, Wesley C. *Measuring Business Cycles.* New York: National Bureau of Economic Research, 1946.

[7] Burns, Arthur F. "The Brookings Inquiry Into Income Distribution and Progress," p. 513. *Q J Econ* 50:476–523 My '45.

Cycles.[8] Schumpeter gives considerable attention to the variability of cycle phases; his last chapter is "The Disappointing Juglar" (the 9-year cycle). Schumpeter's notable idea of *innovations*—which the authors confuse badly with Slutsky's "random shocks" (p. 145)—is of crucial significance here. The authors would find some invaluable suggestions for a pragmatic interpretation of the dynamics of cycles—far beyond the ken of the linguistic labors of Spengler and Sorokin—in Schumpeter's idea of the role of innovations in generating cycles.[9] If the authors would recognize that innovations are like "driving forces" in vibrating systems, or like perturbing forces in periodic orbits, they could serve more effectively the objectives presumably contemplated by the Foundation for the Study of Cycles.

The entire technical apparatus used in this ambitious study is apparently revealed in a footnote on page 174. It seems to comprise merely graphical juggling of deviations from rudimentary moving average "trends"—direct, and rate-of-change. There are ghostly echoes of the "eliminated," "crossed," and "lagged" trends of the "Quadrature Theory" hoax of the middle 1920's (this JOURNAL, March 1924). Naturally the clues to the *mechanics* of cycle generation have eluded the authors. Their avid readers might, perhaps, find some light in their quest for a science of prediction in the better treatises on vibrating systems, e.g., those of Crandall (1926), or Den Hartog (1940). Regrettably, these readers do not appear to have found such light in available recent studies on business cycles.

Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve. *D. J. Finney* (Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford.) Foreword by *F. Tattersfield* (Head of the Department of Insecticides and Fungicides, Rothamsted Experimental Station, Harpenden, Herts, England). London N.W. 1: Cambridge University Press (Bentley House, 200 Euston Road), 1947. Pp. xiii, 256. 18*s*. (New York 11: Macmillan Co. [60 Fifth Ave.]. $3.75.) *Two reviews follow:*

REVIEW BY MARGARET MERRELL
Associate Professor of Biostatistics
School of Hygiene and Public Health, The Johns Hopkins University
Baltimore, Maryland

THE word order of the title of this book rather carries the implication that probit analysis is a method tied to problems of biological response. Although the title may not be important, it is mentioned here because it reflects the presentation of the material. The method of analysis is almost completely identified with the bioassay problem to which it is applied, and vice versa. Thus the book is at the same time a textbook on the problems involved in assaying products through biological responses, and a handbook on the solution to these problems by means of probit analysis. In order to appraise the book, these two aspects will be considered separately.

---

[8] Schumpeter, J A *Business Cycles, Vols 1 and 2*. New York: McGraw-Hill Book Co., 1939.

[9] Cf Goodwin, Richard. "Innovations and Irregularity of Economic Cycles." *R Econ Stat* 28: 95–104 My '46.

The book opens with a definition and discussion of the term biological assay, as "the measurement of the potency of any stimulus . . . by means of the reactions which it produces in living matter" (p. 1). The type of assay most fully treated in the book is that of the all-or-nothing response, such as lived or died. With graded doses of the stimulus, changing proportions of the test animals will respond and a sigmoid response curve usually results. Around this curve, numerous biological and analytical questions arise, concerned with both the variation in the product being tested and the variation in the biological form used for test. Finney gives an excellent discussion of many of these problems. Some of these will be mentioned to give some impression of the scope of the book: the determination of relative potency of two products, and a consideration of situations in which the term "relative potency" has no real meaning; adjustment for natural mortality which occurs independently of the product tested; the factorial design of experiments to avoid the weakness of considering one at a time, a series of factors which may not operate independently; the effect of mixtures of poisons which may be independent and different in their action, similar in their action, or synergistic; the case of a graded rather than an all-or-nothing response. This is a considerably more comprehensive discussion of the problems than is available in other books on bioassay, and will serve a very useful purpose in bringing together material that is scattered rather widely through the literature.

The solution to these problems is oriented to the fitting of the cumulative normal curve by means of the probit transformation. The term "probit" (contraction of "probability unit") was introduced by C. I. Bliss in 1934, for the deviate in a normal distribution with mean 5 and variance unity. His reason for assigning 5 rather than 0 to the mean was a desire to avoid negative numbers.

Finney points out in an historical sketch (pp. 41–47) that the method is much older than the name. He traces it back to 1860 and points out that it has since then been developed anew around problems in a number of different fields. His presentation of the method is, however, so interwoven with the problem to which he applies it that it may be well to state the fundamental steps independent of that problem. These are: (*a*) determining whether a normal curve is suitable for describing the data, and if not, transforming the scale of measurement so that the distribution is normal; (*b*) substituting for the accumulated proportions the corresponding probits, which will then be linearly related to the variate; (*c*) determining the parameters of the straight line relating the probits to the variate; and (*d*) performing goodness of fit tests, and setting up confidence limits about the parameters and the line.

In recent years, J. H. Gaddum, C. I. Bliss, R. A. Fisher, D. J. Finney, and others have contributed extensively to refinements in the methods of estimation, determination of the sampling variations of the constants, and plans of experiments related to this method of analysis.

The book is designed to explain these developments to people dealing with the problem of bioassay. It "is written with the intention of introducing the probit method to many who have previously not ventured to use it, and of presenting some of its more recent developments to those who are already familiar with it" (p. 6). It assumes some but not an extensive knowledge of statistics, and is directed primarily to the applied person. The procedures are explained around arithmetic examples which will be especially useful in acquainting the biologist with the methods. Appendices present a systematic treatment of one example, with all steps clearly described, a mathematical outline of the theory underlying the method, and a series of very useful tables which facilitate the computation.

Probit analysis is first presented in terms of a graphical method of determining the parameters of the straight line, and it is stated that this method is often adequate and serves especially the person who has a long series of estimations to make. The statement (p. 24) that "many experimenters who make use of probit analysis spend time unnecessarily on arithmetic when eye estimation would suffice" is unquestionably true, and it is gratifying to have this point emphasized.

The more exact maximum likelihood estimates obtained by a series of successive approximations are presented in Chapter 4. The reasons stated for using more exact methods are that the observations may be too irregular for any confidence to be placed in an eye estimation of the fitted line, or differences in weights of the observations may be hard to allow for, or the experiment may involve several different factors which are under test which call for more involved analytical techniques. The last reason is a sound one for using the more involved techniques in certain problems, and the second may occasionally be a justification, but the first is dubious. In general it is the very good observations, not the crude ones, that justify elaboration of the analysis. The fact is, that for an enormous number of problems in bioassay, a graphical fitting of the straight line gives adequate estimates; for if the observations are very smooth, the graphical estimates agree well with the maximum likelihood estimates, and if the observations are irregular, the sampling errors are so great that graphical estimations are within sampling range of the more exact estimations, even if they show a sizable arithmetic difference from them. This is true in the examples given by Finney.

The question of graphical versus more exact estimates is intimately connected with the design of experiments, for it involves a balance of gains through time spent on refinements of estimates from a given experiment versus additional experimentation, at various stages of an investigation. As stated earlier, Finney has considerable to say on design of experiments in this field, yet he gives no guidance on this point, saying that "Only experience of the subject and of the experimental technique used can be a sound guide in this matter" (p. 24). This is an issue on which the statistician is frequently consulted, and rightly so, by the experimenter; and the discussion of design of experiment should incorporate consideration of this point.

Methods of analysis of the dosage response curve through other equations than the normal curve are mentioned but not seriously considered. It is stated (p. 47) that none has been found to compare in usefulness with the probit transformation. It is true that none has been as fully explored, and in this sense Finney's statement may be accepted. There is perhaps no reason to expect a book on probit analysis to deal critically with other methods. It is to be hoped, however, that the virtual identification of the problem of bioassay with a particular statistical treatment will not lead the applied man to the view that this is the "correct" analysis to the problem. Referring again to the title, it should be emphasized that this analysis is "*a* statistical treatment" of the problem, and not "*the* statistical treatment." Since it is a widely used treatment at the present time, it is very useful to have this comprehensive summary of the procedures made available, and at the same time, a thorough discussion of the problems in one field where they are applied.

REVIEW BY JOSEPH BERKSON
Division of Biometry and Medical Statistics
Mayo Clinic, Rochester, Minnesota

THIS is a very exceptional if not a unique book. I doubt whether ever before an entire volume has been devoted to the statistical treatment of one particular curve. Confining the specialization still more, only one aspect of the function is considered, for it is not with the normal curve as such that this book deals, but only with its integrated form. Singleness of purpose is further exhibited by the technic of analysis dealt with, for the author gives attention, not to the multitudinous sorts of mathematical methods that conceivably can be applied to the curve, but to one particular type, which is called "probit analysis." Within the probit scheme he largely confines himself to the statistical methods advanced by R. A. Fisher and his students. And finally, the data considered are only those of biologic assay.

Such a concentration of topic and method affords an opportunity for thorough and careful presentation, and this opportunity the author has embraced and exploited effectively. The book as a whole is well planned pedagogically; the simpler elements are expertly separated out for initial presentation, and step by step the more intricate aspects are dealt with in successive chapters. The chapter headings may serve to convey this progress concretely: (1) Introductory; (2) Quantal responses and the dosage-response curve; (3) The estimation of the median effective dose; (4) The maximum likelihood solution; (5) The comparison of effectiveness; (6) Adjustments for natural mortality; (7) Factorial experiments; (8) The toxic action of mixtures of poison; (9) Miscellaneous problems; (10) Graded responses. In Appendix I is given a detailed description, in terms of an example, of the computational procedures; and in Appendix II an account of the mathematical basis of the probit method. There are ample tables, a good list of references, and a good index. I noticed only a few misprints, such as the reference to equation 4.1 on page 55, which apparently refers to an unnumbered equation on page 53; and the year 1940 instead of 1941 given for an article by Garwood. In the list of references the absence of the article by Galton[1], which presented Shepard's calculated table of normal deviates, is conspicuous because it is from this that the probit tables used today are in all likelihood derived.

The basis of the method is the linear transform of the integral of the normal curve, the value of the latter up to X being considered to give the percentage response. The probit Y is obtained by adding 5 to X, in order to avoid negative values in the use of the transform. The transformation makes it possible to represent the relation of the percentage response to the dose by a straight line and to pose the problem of fitting as essentially one of linear regression. The use of linear transformation for curve fitting itself is not novel and is familiar to everyone, at least in the form of fitting an exponential curve by plotting the logarithm. The method was used even for the integrated normal curve long before the initiation of "probits." However, not until Fisher and Bliss devised an appropriate system of weights and "working values," especially the last, was any attempt made to carry over exactly into the transformed scale, the computations that necessarily are defined in terms of the original values.[2]

With a precise linearization thus accomplished, the way was cleared for the analysis of biologic assay, the responses of the latter being considered to be represented by the integrated normal curve, by applying to it all the statistical technics that had been worked out for straight lines. This included in its ramifications almost the entire gamut of methods elaborated by R. A. Fisher, as the list of chapter headings given previously suggests. Since the chapters are comprehensive and clearly written, the volume serves as a good text for the use of these methods themselves, apart from their application to bio-assay.

If one considers it timely to offer a didactic text on probit analysis, as here delimited, to represent bio-assay, then this book is a job well done. However,

[1] Galton, Francis. "Grades and Deviates: Including a Table of Normal Deviates Corresponding to Each Millesimal Grade in the Length of an Array, and a Figure." *Biometrika* 5:400–6 Je '07.

[2] The consideration of weights is related to what I believe is an error in the text. On pages 53–54 we read: "This quantity is $Snw(y-\bar{y})^2$ reduced by $[Snw(x-\bar{x})(y-\bar{y})]^2/Snw(x-\bar{x})^2$; the calculation is shown in Table 6, and $\chi^2=1.67$, which differs from the χ^2 of Table 7 only on account of rounding-off errors." The value of Table 7 referred to, 1.62, is the actual χ^2 of the observations in relation to the fitted curve defined as $\sum\frac{(o-t)^2}{t}$. In the probit iteration, at each step, we use, for the regression being determined, approximate weights and working values derived from the previously determined regression. It is the sum of the squared residuals of the working probits around the fitted line thus falsely weighted that is given by the expression referred to The χ^2 directly calculated refers to the actual observations and is equal to the sum of their squared residuals with the weights of the present solution. These two sums will not be equal, no matter how many decimal places are carried, until finally when the maximum likelihood solution is obtained. In fact a signal, and a useful one, as to whether this point has been reached is just this comparison.

I myself doubt the appropriateness of such a volume. That the probit is a handy and generally good-enough tool for dealing statistically with the data of biologic assay need not be doubted. If easily applied methods for fitting are utilized and if the associated formulas are considered as suggestive only and are used with ample safety factors, no harm, and very probably much good, can come of its use. But the impression often conveyed by its devotees that this particular transform is a thing unique, possessing some intrinsic superiority, and that the methods used are mathematically "exact" and therefor unexceptionable, is open to question.

Out of some 200 pages of text, the author devotes only about one to curves other than the integrated normal curve and says: "None of these relationships has yet been found to compare in usefulness with the probit transformation." It would be more accurate to say that no comprehensive examination has been attempted of the comparative merits of different transforms and that such comparisons as have been made have not been examined adequately if at all by the author of the present text. Some studies indicate that there are certain advantages in other transforms. The angular transformation, for instance, has been found simpler to fit[3] and the logistic not only easier to fit but better in accord with bio-assay data.[4] Before a particular curve is adopted as the statistical norm, do we not need to know whether it represents the actual physicochemical mechanism of the bio-assay reaction; that is, do we not need a study of the differential equations representing the theoretical mechanisms involved? As a minimal responsibility it would seem that the author of so definitive a text should himself have ascertained which of several proposed curves *fits the data best*, before committing a public that is gullible with regard to any statement printed in a textbook to the curve that happened to have emphasis in his statistical schooling. In fairness, however, it should be said that the methods worked out for the probit transform can, with appropriate modification, often be applied to other linear transforms, should it be desired. A statement to that effect would, I believe, strengthen the present work.

The sectarianism shown by the author in regard to the choice of curve is exhibited likewise in respect of the statistical methods advanced for applying it. The method advocated for fitting the curve is that of maximum likelihood, as proposed by R. A. Fisher, rather than least squares, on which most of us oldsters were brought up. For the situation of biologic assay as dealt with here these two methods do not give the same estimates, as the author clearly recognizes (p. 54).[5] An example for which I have made some necessary calculations is one used by Fisher and Yates[6] and for which Garwood[7] calculated the maximum likelihood solution by the probit method. Four successive approximations were required before constancy of the estimates of the parameters was achieved. For the first provisional solution the estimates of a and b were respectively 4.600 and 0.600; for the fourth solution they were 4.566 and 0.713, showing an appreciable difference between the first and last estimates. Now the likelihood was highest for the last solution, but the successive values of χ^2 beginning with the first provisional solution were 3.54, 3.67, 3.77 and 3.79. The χ^2 test is often used as a measure of goodness of fit, and its use is in fact a part of the prescribed procedure of probit analysis. By this criterion, in the successive iterations in this example we were laboriously moving toward a worse and worse solution! Which solution is the "best": the first, of which the value of χ^2 was smallest; or the last, for which the likelihood was largest? Which of the two estimates has the smaller variance?[8] And does not the use of the χ^2 test require that no part of any possible excess of χ^2 found be due to the method of estimate used?

---

[3] Knudsen, Lila F., and Curtis, Jack M. "The Use of the Angular Transformation in Biological Assay." *J Am Stat Assn* 42:282–96 Je '47.

[4] Berkson, Joseph. "Application of the Logistic Function to Bio-Assay." *J Am Stat Assn* 39: 357–65 S '44.

[5] This apparently has not always been clear to writers on the subject. Thus we find Irwin and Cheeseman (*Sup J Royal Stat Soc* 6:174–85 '39) saying in an article which is often referred to: " . . . as is to be expected from the efficiency of the maximum likelihood solution, which in this case is the ordinary least square fit."

These are not impertinent questions; indeed they must be answered before we can say that the basic procedure used in probit analysis rests on a logical foundation. For the difference between graphic fits by eye and the exact maximum likelihood solution is of the same order as the difference between minimum χ^2 and maximum likelihood estimates. If we are to proceed from fitting by eye, using a rather laborious arithmetic procedure, toward a more refined solution, it can only be because we are assuredly moving toward a better solution. But if, as is possible, the minimum χ^2 solution is the better, then in seeking a maximum likelihood solution we are, as would be illustrated in the case of Garwood, actually achieving a worse estimate than the one with which we started.

My own questions regarding the probit method for fitting are for the most part on an intuitive and "common sense" basis and are little assisted or embarrassed by knowledge of the higher mathematics. However, misgivings evidently have been felt by the mathematically mature and competent. Thus, speaking of the probit solution, Thompson[9] said: "A general theorem of convergence has not been proved for any set of points (log. *Di*, *pi*), or with certain specific exceptions, where we are required to fit a given curve or a straight line in transformed coordinates in the indicated manner. Indeed, the general theorem without such exceptions can easily be disproved . . . and reference to it as 'the exact' method . . . may lead to a false appreciation."

---

[6] Fisher, Ronald A., and Yates, Frank. *Statistical Tables for Biological, Agricultural, and Medical Research*, Second Edition. Edinburgh, Scotland: Oliver & Boyd, Ltd., 1943.

[7] Garwood, F. "The Application of Maximum Likelihood to Dosage-Mortality Curves." *Biometrika* 32:46–58 Ja '41.

[8] To many the question of which has the smaller variance will appear to be authoritatively answered by Finney in the following quotation (p. 209): "Fisher (1922, 1925) has shown that estimates of the parameters which maximize the likelihood are *efficient* in the sense of having minimal sampling variance in large samples." Since for the case in discussion the maximum likelihood and minimum χ^2 estimates are different, it would appear that either both their variances are equal or the variance of one is smaller and that the one given by maximum likelihood. Inquiry discloses, however, that such simple interpretation reveals a lack of insight into the delphic implications of English when spoken by mathematicians. For it appears that minimum χ^2 estimates are also "efficient" estimates and apparently it is not in contradiction with the quoted dictum that they may give for samples of say 1,000, estimates with smaller variance than those of maximum likelihood.

[9] Thompson, William R. "Use of Moving Averages and Interpolation to Estimate Median-Effective Dose." *Bacteriol R* 11:115–45 Je '47.

Similar doubts will assail one on careful examination of many other phases of the probit analysis. On the basis of practical experience I have always felt instinctively that the use of observed values of zero or 100 per cent when these are asymptotic values is hazardous in curve fitting procedures. Others apparently hold similar views; for instance, Knudsen and Curtis in advocating the arc-sine transform advise against the use of any observations outside the range of 0.05–0.95. Cannot the basis of this hesitancy to use the absolute rates be found in mathematical considerations? In assuming that the integrated normal curve represents the true quantal responses to increased dosage, it is implied that for very large samples, a rate of zero or 100 per cent will not occur except with infinite doses. This is of course in marked contradiction with actual observations, in which such rates are always encountered with finite and even moderate doses. In his basic article on "The case of zero survivors" Fisher[10] does not evidently make any allowance for this difference between theory and fact. But it is not to be taken for granted that this difference between the mathematical model and reality is without practical consequence, and that it may be disregarded without biasing the estimates. In view of the practical worker's hesitation to use absolute rates, the burden of proof would appear to rest with those advocating their inclusion in statistical schemes. Do we not have the right to require of the mathematical statistician that he demonstrate that the consequences of likely departures from theory are negligible, before developments based on the theory are considered definitely established?

The fallacious practice of applying a test of significance by first offering lip service to the doctrine that a nonsignificant P (not low, that is) does not establish the hypothesis tested but only fails to disprove it, and then *acting* as if it did establish it, a fallacy that has been warned against repeatedly by authorities in mathematical statistics,[11] is illustrated in Finney's book in the use of parallelism of probit lines for estimating relative potency. First parallelism is *assumed* in order to make the estimates, then a χ^2 test is performed for parallelism on the basis for instance of two degrees of freedom (p. 71). *Mirabile dictu,* the P turns out not significantly small. *So the estimates based on the assumption of parallelism are accepted.* Before such a test can be considered to rest on a sound mathematical foundation we must have some knowledge of the *power* of the test relative to the alternative of nonparallelism. How large a departure from parallelism must exist, I wonder, before the

[10] Fisher, R. A. "The Case of Zero Survivors." An appendix to "The Calculation of the Dosage-Mortality Curve," pp. 136–67, by C. I. Bliss. *Ann Appl Biol* 22:164–5 F '35.

[11] *See,* for instance, A. Wald's review of M. G. Kendall's *The Advanced Theory of Statistics,* Vol. II in *J Am Stat Assn* 42:185–6 Mr '47.

χ^2 test, as used here on the basis of two degrees of freedom, will probably show a significantly small P? And is it not a fact that departure from parallelism much smaller than this will render seriously inaccurate estimates of the parameters which have been made on the basis of an assumed parallelism? Perhaps Finney will have this calculation for us in the next edition of his book. Or perhaps he will find that parallelism cannot be tested effectively by means of tests of significance based on two degrees of freedom.

Another matter inviting attention is the tacit assumption made throughout the text that there is no error in x, the measure of dosage. Actually this is an idealization, for to err is human, even in the direction of the abscissa. Do we not need a mathematical investigation to ascertain the possible effect of such errors? So far as I know, no such inquiries have been made, but it is not hard to see intuitively that such errors, even if moderate in size, may play havoc with many of the formulations set forth as "exact" in this book.

Aside from such specific questions, I sense a broad uncertainty as to whether the structure of assumptions as a whole underlying the formulas is warranted for the data as they actually are encountered. It is notable that this entire book contains not a single repetition of experiment to test whether the theoretical error formulas represent the real variations which occur in repeated trials. My own practical experience, admittedly small, indicates that the formulas underestimate the actual errors considerably and sometimes vastly; and this conclusion is confirmed by veteran workers in the field of bioassay. Sometimes this sort of objection is met with the rejoinder that the formulas give only the "statistical" or "sampling error"; but, if dosage to patients is guided by measures of error as given by these formulas and a patient succumbs from an overdose, it will be little consolation to reflect that it was not the sampling error that killed him.

For the reasons cited and other similar ones we cannot be certain that probit analysis furnishes the exact method for bioassay nor even that, of the methods proposed, it necessarily is the best. At the moment about all we can be sure of is that it is the most elaborate.

Manual de Estatistica. (Manual of Statistics.) *Amaro D. Guerreiro.* Instituto Nacional de Estatistica, Portugal. Lisboa; Tipografia Matemática, Lda., 1947. Pp. 340.

REVIEW BY HUGO MUENCH
Professor of Biostatistics, School of Public Health
Harvard University, Boston, Massachusetts

THE author states that his publication was prompted by the limitations of the study of statistics in Portugal, and the scarcity of statistical bibliography in Portuguese. He is careful to define the book not as a comprehensive text of statistics but rather as a supplement to courses given by statistics departments.

Given a book of this type, it is probably too much to expect a critical appraisal of the methods described since it is in the nature of a laboratory manual or of a mnemonic aid in the performance of calculating procedures. Even so, it is disappointing to encounter a wide range of statistical methodology without presentation of rationale or caution as to conditions of legitimate use. It is even more disappointing to find, in the wealth of illustrative material, methods applied to conditions not fulfilling the underlying assumptions under which such methods are justified.

Normal correlation is, of course (and properly, within due limits of caution) applied to material whose distribution is not strictly normal. From the pedagogic viewpoint, however, it seems to this reviewer that illustrative material should conform as closely as possible to prescribed conditions. It is therefore something of a shock to find methods of normal correlation and of partial and multiple correlation extensively illustrated by material which is inherently nonnormal. Incidentally, Pearson's "eta" is presented as a measure of "curvilinear correlation."

To begin with, the author's presentation of distributions seems confused. Concepts of meso-, lepto- and platykurtosis are illustrated by a graph of three distributions, all apparently normal but with differing scatter. Later on, relationships between quartiles and standard deviations are stated as valid for "symmetrical" distributions rather than for normal, and a series of measures of asymmetry and kurtosis is presented without discussion.

The lack of critical presentation is most felt, perhaps, in the long chapter devoted to index numbers, though its absence is felt also in the two final chapters devoted to curve fitting. The first of these, entitled "Analysis of Time Series" is given to methods of fitting curves by differences (the strange statement is made that "if first differences are normal, the logistic should be used"); by least squares (parabolas to the third order); by selected points (logistic); and by least squares of logarithms (exponential and hyperbola). Smoothing by the use of moving averages is discussed; cyclic functions are nowhere mentioned.

The final chapter presents the fitting of various frequency functions. Binomial distributions are somewhat vaguely linked to the normal curve which, for some reason, is transformed so that its maximum ordinate is set equal to one, rather than its area. The Pearson system of curves is not explained, but Elderton's table of these functions and their characteristics is included and a Type I curve is fitted at some length.

Three appendices present tables of chi-square, of functions of the normal curve (ordinate $=1$ when $x=0$) and of the log Gamma function. There is an extensive table of contents, as well as indexes of tables and of graphs, but there is no alphabetic index. The table of errata fails to mention at least one mistake: the writing of the normal curve function on page 240.

Mr. Guerreiro's book should be recommended only to such students as have already acquired a good background in the use of statistical method and who therefore would be able to select, from the material here presented, procedures applicable to the nature and the limitations of their data. Taken by itself it would in no way prepare a student to follow a book such, for example, as Jorge Kingston's *A Teoria da Indução Estatística.*

An Outline of Statistics, Third Edition. *Samuel Hays* (Formerly Lecturer in Commercial Subjects, Doncaster Technical College, Doncaster, Yorks, England). London S.W. 19: Longmans Green and Co. Ltd. (43 Albert Drive), 1947, Pp. vii, 254. 8s. 6*d.* (New York 3: Longmans Green and Co. [55 Fifth Ave.]. $2.25.) *Two reviews follow:*

REVIEW BY HAROLD NISSELSON
Statistician, Bureau of Census, Washington, D. C.

THIS is the third edition of an English primer written for "Commercial Students in our Technical Colleges" and the "general reader of economic literature." To this end, it is intended to cover "the requirements of the National Certificate in Commerce and the various Accountancy examinations."

Roughly two-thirds of the text is devoted to a nontechnical discussion, in the traditional spirit, of descriptive methods in statistics. Of the remainder, half is given over to a consideration of published sources of statistics on England and the United Kingdom, and the balance to a set of routine exercises (with answers). The principal revisions consist of a new chapter dealing with rounding errors, an appendix on quality control, and the inclusion of postwar sources of data. Since the book is almost pocket-size, the amount of space that could be devoted to any one topic is fairly limited. This is reflected in the progressively sketchier treatment of later topics; in particular, time series, correlation and index numbers. However, the style is much more agreeable than this brevity would indicate.

The remaining comments apply to the book as a text in statistics, rather than its adequacy for the examinations cited. The chapters on official and other published sources of data for England, and an additional one on statistical measures (unit costs, turnover, etc.) in common business use, are the most valuable sections of the book and should prove very useful. By contrast, the technical material is far less satisfactory. In general, the approach to statistics is one of merely summarizing a set of data. Hence, it is considered a limitation of the arithmetic mean that for discrete variables (e.g., number of children per family) it may not "represent an actual item at all"—the mean being interpreted as an observation rather than as a parameter of the population. There is almost no mention of the role of sampling variation in interpreting statistical data, and the choice of the few instances where such reference is made is quite unexpected. Thus, a formula is given for the probable error of a mean or total due to rounding of the individual figures involved, but there is no mention of the probable error of a sample mean or total as an estimate of the corresponding quantity in the population sampled. (The probable error due to rounding as given is incorrect, and the statement is also made that the probable error is "the most likely value of the error to occur.") Again, a two-page appendix explains the probability that a population proportion lies in a specified range based on the proportion observed from a sample; and an unsatisfactory cookbook rule is given for judging the significance of an observed correlation. The examples cited here are typical of a number of misstatements and unguarded remarks on methodological material which appear at all levels of sophistication.

For teaching purposes the unconventional terminology is awkward: for example, "coefficient of dispersion" for the coefficient of variation, "ratio of variation" for the regression coefficient, definition of the first moment of a set of numbers as the mean deviation and the second moment as the standard deviation. Moreover, even for a one-semester introductory course, extensive supplementation by lectures would be necessary.

In future editions, the usefulness of this book could be greatly enhanced by pruning or eliminating entirely the chapter on rounding errors, and, if necessary, issuing the exercises separately. The space gained in this way could then be devoted to developing and illustrating the basic ideas of statistical inference, and the relations between a sample and the population on which it is to provide information.

REVIEW BY EDGAR Z. PALMER
Professor of Statistics and Director of Bureau of Business Research
University of Nebraska, Lincoln, Nebraska

THE preface to this little book states that it "is intended primarily for Commercial Students in our Technical Colleges and for students interested in professional examinations." It may be described as British in form and somewhat American in approach. The total amount of material included is scanty compared with most textbooks. The usual topics of collection, tabulation, and "diagrammatic" representation are touched upon, as well as averages, dispersion, index numbers, and correlation. In the chapter on time series most of the attention is given to the seasonal index.

The book gives evidence of careless preparation. It is stated that a death rate of 12.5 means a value anywhere between 12 and 13 (p. 22); a frequency polygon is smoothed by the use of a weird collection of curves joining the tops of the bars (p. 62); the third moment measure for skewness is given in terms of the cubes of the deviations, without indication of summation (p. 98); moving averages, plotted at the *ends* of their periods, are supposed to illustrate the smoothing process (p. 106); a square root sign appears without a figure under it (p. 205). Formulas are a crude combination of words and symbols, and fractions are mixed with decimals in a confusing manner. Sampling and quality control have been relegated to two appendices, the former topic including only the sampling error of a proportion, and the latter introducing for the first and only time the sampling error of the mean.

It is probable that for the quality of student for which this is intended it is quite teachable. The writing is clear, and illustrations of computations, tables, and graphs are generous. The only interest for American readers is in five short chapters on the sources of British statistics, including one chapter each on vital, employment, overseas trade, and price data.

Market and Marketing Analysis. *Myron S. Heidingsfield* (Assistant Professor of Marketing. Temple University, Philadelphia 22, Pa.) and *Albert B. Blankenship* (Managing Director, National Analysts, Inc., 15th and Locust Streets, Philadelphia 2, Pa.). New York 10: Henry Holt and Co., Inc. (257 Fourth Ave.), 1947. Pp. x, 335. $3.00. *Two reviews follow:*

REVIEW BY LESTER R. FRANKEL

Statistician, Marketing and Research Division
Dun & Bradstreet, Inc., 290 Broadway, New York 8, N. Y.

THE use and application of marketing research to the solution of business problems has received general acceptance during the past few years. Since many of the procedures used in this field are based upon somewhat specialized statistical techniques, a warm welcome would greet a book, such as this one, whose purpose is to elucidate some of these methods.

In spite of the excellent choice of topics, the orderly presentation of subjects and the simplicity of exposition, this book, as it now stands, fails to serve the above purpose. This results from the fact that the text contains an unduly high proportion of mistakes, misinterpretations, and confusing statements. Some of the errors can be detected by the bright student or astute businessman. Other statements are inconsistent with correct statistical usage. Businessmen with some knowledge of accounting procedures will find some of the material confusing. Students of marketing who read the book and then decide to specialize in statistics will have to unlearn many of the techniques which are incorrectly described here.

The material covered in the text is presented in a logical and systematic order within the framework of five parts. The first part, consisting of three chapters, is introductory in nature. It discusses the need of research to the efficient operation and planning of a business concern and points out that sources of information are available from both internal and external sources. Internal analysis, it is explained, makes use of current and past operating records of the concern. Such records are usually in the form of balance sheets and profit and loss statements. In external analysis use is made of surveys, statistical series and other sources outside of the concern.

The second part deals with the use of internal analysis in the solution of marketing problems. Although the study of the operations, problems, analysis, and interpretation of facts pertaining to business concerns is somewhat of a science today, the two chapters in this part, "Analysis of Business Records" and "The Use of Standards in Internal Analysis," use business terms and expressions quite loosely. The effect is confusing, whether the looseness is due to careless wording or to a lack of detailed knowledge of a very exacting subject.

On page 41, for example, it is said that ratios measure "degree of solvency" and "percentage of borrowed assets." "Degree of solvency" is not an established term in business or the economics of business, and the text does not make it clear whether it means that a particular business enterprise is 50 per cent solvent, 70 per cent solvent, 90 per cent solvent, or exactly what. The text goes on to state that "degree of solvency" can be measured by the ratio of "assets versus liabilities." Here we become still more confused, as we are not told what assets or what liabilities are to be included. We learn a little later that what is meant is current assets and current liabilities, although on page 55, where the ratios are listed in fraction form, the fraction again is based on the very indefinite terms of "assets" and "liabilities." The ratio of current assets to current liabilities is no measure of such indefinite ideas as "degree of solvency" or "percentage of borrowed assets." It is really a measure of the ability of a concern to meet its immediate liabilities on scheduled time, and even as such a measure, interpretation must be made with extreme care and from fairly wide knowledge or experience. The second expression, "percentage of borrowed assets," has no meaning. No business concern borrows assets. Assets are purchased, and until the purchase price is paid, the liability for the purchase is on the books. The purchase might be raw material to be paid in 10 days, or real estate and building to be paid gradually over a period of 10 years.

In several of the ratios which are mentioned, "net worth" is used as a component. Net worth is the dollar difference between all assets and all liabilities. Occasionally a concern has total assets of $1,200,000 of which $1,000,000 is an intangible item such as goodwill, trade marks, or formulae. If the $1,000,000 is deducted, we arrive at a term of "tangible net worth" which should be used in these compilations. Otherwise homogeneity, which is described as an absolute essential on page 64, is completely lacking. This is a matter of very considerable importance and no explanation of the difference between net worth and tangible net worth appears in the text. As a matter of fact, the term "net worth" is used exclusively in the text.

The material presented in Part 3 provides a very good summary of the use of sample surveys in marketing research. In many respects this part is the best in the book, especially from the viewpoints of clarity and accuracy. But, even here, we find some errors. For example, sampling is first introduced in the chapter on "Planning the Survey Operation." At the beginning of this chapter (p. 83) we find the statement that, in sample surveys, "The important thing here is to make absolutely certain that those selected to represent the total population have all the social, economic, and other characteristics which typify the original group." A few lines later we find the statement, "Statistically, this is usually secured by what is called a random selection." Of course, in random selection we are *never* absolutely certain that the sample has all of the characteristics of the population. In random selection, we rely upon the statistical law of large numbers to approximate this objective.

The issue of area versus quota sampling is discussed in the same part. It was gratifying to find that through an objective analysis of the two methods the authors succeed in conveying to the readers the fallacies of quota sampling. However, it is unfortunate that the authors persist in using the term "stratified sampling" synonymously with "quota sampling."

In the remaining chapters of this part we find excellent treatments of the problems of questionnaire development, and of the collection and tabulation of the data. We are given the benefit of the authors' wide experience in this field.

"Special Applications of Research Technique" is the title of Part 4, which includes a description and evaluation of the methods of establishing sales potentials and quotas, and a discussion of some of the problems of radio research. It is interesting to note that, in commenting upon the use of correlation analysis in determining market indexes, the authors state (p. 229), "This method has the disadvantage of requiring a trained statistician."

The part of the book dealing with "Technical Statistical Procedures" is the part containing most of the errors. Some of the errors are due to carelessness and others are due to the authors' eagerness to describe techniques which they evidently have never used. The first chapter in this section describes fundamental statistical techniques, material which is found in all elementary texts on statistics, and yet there are a number of errors. For example, in describing the methods of computing the arithmetic mean and median, a series of 33 numbers is grouped in the form of a frequency table. In computing the mean the midpoint of the class interval, say 13 to 15, is taken at 14. However, in calculating the median we find the statement, "The actual range of the scores represented by the 13–15 class interval includes those that are barely above 12 and those extending up as high as 15 but no higher." The class interval has shifted one half of a unit. However, the correct answer is arrived at in this particular instance through the commission of another, but compensating, error. The position of the median item is computed by taking one half of the total frequency plus one instead of one half of the total frequency.

There is confusion in the distinction between the standard deviation and the standard error. It is stated on page 273 that "The standard error, numerically, means just about the same thing as the standard deviation." If 50 per cent of the public has done a certain thing "there is no way to compute a standard deviation of such a figure." The only thing one can do, say the authors, is to get the standard error which is given by the familiar $\sqrt{pq/n}$ formula. The distinction between statistical description and statistical inference is not recognized.

The treatment of "Sample Size and Accuracy" is somewhat puzzling. It is emphasized at the outset that statistical formulas for sampling errors can be applied only when the sample has been drawn on a random basis. We are warned that the formulas do not apply in quota sampling and that modifications of the random sampling formulas have to be made when stratification or area sampling has been employed. However, the authors state that once data have been collected, regardless of the particular sampling used, the sampling error can be computed by established techniques! They have evidently fallen into this fallacy because of their ignorance of the difference between the standard deviation and the standard error.

In a table on page 280 illustrating the cut-off method for determining the size of an adequate sample we find that among 50 interviews 49.2 per cent answered "yes" to a particular question and that among 100 interviews 50.1 per cent answered "yes." Since such results are arithmetically impossible when each interview represents a single individual, we are led to believe that an interview may represent several individuals However, there is no indication in the text that this is the case.

On page 290 the following statement appears: "The midpoint between the perfect positive and perfect negative correlation is indicated statistically by a correlation of .00. This means that there is no relationship whatever between the variables being analyzed." What should have been said is that there is no simple linear relationship between the variables.

It is stated on page 295 that, "Mathematically, it is seldom sound to . . . [compute a linear correlation coefficient] with fewer than 100 cases." Provided the correct test of significance is used, it is sound to compute r with as few as three cases.

On page 305 in discussing the uses of the chi-square test it is suggested that, "It might be used to determine the significance of sales estimates (as incorporated in the sales budget) against sales returns." The comparison between an actual series and a theoretical series through the use of chi-square can be made only when both series consist of frequencies.

It is unfortunate that this book did not receive a thorough review prior to its publication. The layout, presentation, and style are above par for a text of this type. However, because of the great number of errors present, it is not very well suited for classroom or general use. It is hoped that a revised edition of the book will soon appear.

REVIEW BY PHILIP J. MCCARTHY

Assistant Professor of Sociology, Cornell University
Ithaca, New York

THE topics considered in this book fall into three categories: the use of internal company records for market research, the application of survey techniques for obtaining data external to the company, and the statistical concepts which are frequently required in considering either or both of the other two topics. All three are well illustrated with practical examples. The Preface states that the book has been prepared with intentional simplicity for students in marketing research and for businessmen who buy or otherwise make use of commercial research. This, in itself, is a desirable characteristic of a book for beginners. However, when considered in conjunction with the fact that much of the material on survey techniques has already been presented in a simple and more exhaustive manner by one of the authors,[1] and with the fact that statistical techniques are very difficult to treat in a simple manner, this simplicity becomes the greatest shortcoming of the book.

The discussion of internal company records as a tool for research is an innovation for a marketing research book. As such, it should prove particularly interesting and valuable for those individuals who have entered the field of commercial research through academic disciplines other than economics or business administration. However, it is doubtful whether the student of marketing or the businessman will find this approach new. Most students of marketing will already have covered the field of internal record analysis in more detail in their courses in business statistics and accounting, and it seems likely that the businessman would have long since been forced into a careful study of the internal records of his company.

Survey techniques are skimmed over lightly, the presentation being very similar to and less detailed than that of the previously noted book by Blankenship. The authors do maintain a critical attitude throughout this section, and their comments on methods and techniques are unusually well chosen and to the point. They have perhaps leaned over backwards to stress the use of random methods in the design of sampling procedures, but considering the widespread use of other methods in marketing research, this can certainly do nothing but good. They do attempt to wrap all sampling methods up into four neat packages (an impossible task), and they do say nothing about how the businessman or the commercial research organization should determine the limits of error which can be tolerated in any given survey operation. Unfortunately there is very little information on this topic, and some considered thought needs to be devoted to it.

The final topic covered in the book, namely statistical techniques, is another attempt to convey statistical understanding to persons without any mathematical background. The present treatment is not distinguished in any way, and when one considers that measures of central tendency, measures of dispersion, sample size and accuracy, correlation, analysis of variance, analysis of covariance, factor analysis, chi-square and sequential analysis are all touched upon in the short space of fifty pages, there is serious reason to doubt whether this section will serve any useful purpose. The interested reader might better be referred to some such book as *How to Read Statistics* by R. L. C. Butsch (*see* reviews in this JOURNAL, December 1946).

O Custo de Produção do Homem Adulto e Sua Variação em Relação à Mortalidade. (The Cost of Production of the Human Adult and Its Variation in Relation to Mortality.) *Giorgio Mortara* (Technical Advisor, National Census Commission, Brazilian Institute of Geography and Statistics, Rio de Janeiro, Brazil). Estudos Brasileiros de Demografia, Monografia No. 2. Rio de Janeiro, Brazil: Livraria Kosmos Editôra, 1946. Pp. 152, 40 cruzeiros. Paper.

---

[1] Blankenship, Albert B *Consumer and Opinion Research*, New York: Harper & Brothers, 1943. Pp. xi, 238.

REVIEW BY T. N. E. GREVILLE
Principal Actuarial Mathematician, National Office of Vital Statistics
U. S. Public Health Service, Washington, D. C.

THIS is the second of a series of monographs entitled "Brazilian Demographic Studies," being published under the auspices of the Getúlio Vargas Foundation. Pareto in 1893 (Giornale degli Economisti, 1893, 2nd half-year, pp. 451–456) discussed the cost of production of a human adult, defined as the total cost of upbringing of an entire generation from birth to age 15 or 20 (including the children who fail to survive) divided by the number of survivors at the selected age. Finding this cost to be only slightly less in countries with low infant mortality than in those having high infant death rates, Pareto erroneously concluded that a low infant mortality tends to be offset by a high death rate in later childhood and adolescence. The author of the present monograph conclusively disproves this contention, marshaling a mass of data derived from life tables for a number of countries at various periods to show that a reduction in infant mortality has invariably been associated with a corresponding reduction at subsequent ages.

Postulating a scale of relative living expenses by age increasing in geometric progression from 26.79 in the first year of life to 100.0 for the twentieth and each subsequent year, he obtains values, in terms of the conventional units so defined, for the total cost of production (up to age 15) of a human adult in various countries and epochs, including the proportionate share of the cost occasioned by those who died prior to age 15. Subdividing the result into two parts, corresponding to the proportionate share of the cost of upbringing of the survivors and nonsurvivors separately, he shows that even in those countries having the highest infant mortality, the latter portion of the cost constitutes only slightly more than ten per cent of the total. This figure represents, therefore, the maximum reduction in the total cost that could, in theory, be effected in a given country by reducing the death rate in infancy and childhood. However, if one considers separately the percentage change in the portion of the cost which relates to the nonsurvivors, there has been a spectacular reduction. For example, it was 64 per cent for white males in the United States between 1900–1902 and 1939.

On the basis of available data, and making certain assumptions, the author attempts to illustrate his results in terms of concrete monetary values, estimating the average cost of production of a human adult under the economic conditions existing in Brazil in 1939 to have been between 8,700 and 8,750 cruzeiros. At the rate of exchange then prevailing this would be equivalent to approximately $450 in United States money.

The style is lucid, and the argument convincing. In general, the conclusions stated appear to be amply supported; in fact, the documentation, in some instances, is perhaps unnecessarily full, as some of the main conclusions are today hardly open to dispute. The author promises that a subsequent monograph of the series will deal with the average productivity of an adult during his lifetime and its variation in relation to mortality.

Forecasting Sales. *G. Clark Thompson.* Conference Board Reports, Studies in Business Policy, No. 25. New York 17: National Industrial Conference Board, Inc. (247 Park Ave.), 1947. Pp. 47. Paper. Price on application. *Two reviews follow:*

REVIEW BY CHARLES F. ROOS
President, The Econometric Institute, Inc.
500 Fifth Ave., New York 18, N. Y.

NO LONGER than ten years ago sales forecasting by most companies was confined to tabulating the guesses or hopes of the sales department; as late as 1940 only a few companies recognized that their sales are closely geared to national or regional purchasing power and other economic variables that can be satisfactorily forecast. As the Foreword of *Forecasting Sales* states,

> Until recent times, most companies have contented themselves with crude appraisals of the general business outlook. The gradual development of economic and marketing research techniques, however, have provided businessmen with an opportunity to make more accurate sales forecasts than had previously been possible. Today, numerous companies regularly forecast annual sales within 10% of actual performance and often come within 1%.
>
> A realization of the value of such forecasts, together with increased confidence in present-day forecasting techniques, has encouraged many companies to attempt to improve their methods or to develop new ones.

The main purpose of this report is to illustrate forecasting techniques currently being used by companies in such diverse fields as food processing, metal fabricating and manufacturing, air conditioning equipment, furniture, electrical equipment, chemicals, rubber products and building materials. The methods illustrated for forecasting sales include the polling of executive and sales opinions, correlation analysis, trends and cycles and the use of mathematical formulas which last in reality involve the use of statistical complexes and accounting techniques. Actual examples are used to develop the techniques and, consequently, they are characterized by uneven degrees of statistical refinement and imagination.

The first example presented under the section devoted to statistical methods is that of a metals company. Joint study by outside consultants and a company specialist revealed a definite relation between the sales of the company's products and various special groupings of industries included in the Federal Reserve Board index of industrial production. The example presented correlates the unit sales of Product "A" with the Federal Reserve Board index of industrial production and the ratio of automobile production to this index. Thus, the approach avoids the pitfall of using correlated variables in a multiple correlation. The "fit" between actual and estimated sales is quite good, as it should be, because the product is sold to many industries but in volume to the automobile industry. This means that if the company can secure reliable forecasts of the FRB index of industrial production and of automobile production it can make reliable forecasts of the sale of Product "A." The company obtains these general forecasts from a private research organization which specializes in this work.

Another example compares the plumbing sales of a manufacturer with construction contracts. The agreement, of course, is generally good. However, there are important deviations which seem too large to be ignored. The reviewer has found that these plumbing sales can be explained more satisfactorily in terms of Supernumerary Income (disposable income less cost of subsistence living) and construction contracts. The Supernumerary Income is the purchasing power available for purchases of plumbing supplies for modernization. This income is sometimes called Discretionary Income.

The projection of trends is the method used by an electric power company. It projects the trends of (1) population, (2) meters or customers, (3) electric power consumption per customer, and (4) revenues. The trend of each is studied separately and also in relation to the other three. As crosschecks or links the company also studies (*a*) the meter ratio (number of people per meter), (*b*) sales per meter, (*c*) revenue per meter, (*d*) revenue per unit, and (*e*) revenue per unit vs. consumption per meter. With the aid of mathematical formulas and a certain amount of judgment the trends of the link factors are projected.

Industry replacement sales of automobile tires are forecast by a mathematical formula which includes the number of automobiles over two years of age, consumption of gasoline per car of these cars, a factor measuring tire quality, and sales of camelback, the unvulcanized rubber used for recapping. A curve can be readily constructed from life tables to show the number of cars over two years old, the tire quality is known, and sales of camelback are not large enough to introduce much error. Consequently, the principal problem is that of forecasting gasoline consumption. For this purpose the company projects the trend of gasoline consumption per car by mathematical extrapolation modified to allow for special conditions which will, in the forecaster's judgment, affect the demand in future years. Much more accurate forecasts of gasoline consumption per car can, however, be made by relating this variable to disposable income and gasoline price and forecasting disposable income. Gasoline price is related to this income, taxes, and refinery costs.

Some companies use several methods for forecasting sales and then reconcile the forecasts in meetings of executive or planning committees. In such instances, the forecast has broad authority and acceptance throughout the organization. While it is recognized that the forecast may still be wrong, it is accepted as the best estimate the organization can make based on economic and statistical information as checked by the business experience of the company's top officers.

It is now possible for companies to take the next step and plan sales budgets and allocations by counties. The Office Equipment Manufacturers' Institute has arranged with the Social Security Board to tabulate for each county in the United States employment, payrolls and the number of establishments broken down to size groups. These will be provided for the main classifications of mining, manufacturing, wholesale trade, law offices, medical and health services, and nonprofit organizations. It will now be possible for companies to make studies of how their sales gear to these county data and then by the use of tabulating machines to set realistic sales budgets by counties and sales territories.

This trend toward better forecasting, in the reviewer's opinion, will probably lessen the intensity of business cycles and will certainly lead to the dampening or elimination of seasonals. This is especially important at a time when union leaders are beginning to press for guaranteed annual wages.

REVIEW BY H. D. WOLFE

Director, Market Research Department
Colgate-Palmolive-Peet Co., Jersey City, N. J.

THIS small pamphlet on methods and practices of forecasting sales is one of the most helpful and illuminating contributions ever made on this most vital subject. Much has been written about distribution costs and the fact that they may or may not be too high. Thompson, through exposition of method buttressed by actual cases of the practices employed by fifteen large industrial companies, demonstrates how accurate sales forecasts can lower distribution costs. Costs of production, selling prices, and profits are, of course, dependent on sales volume. The sales forecast of volume (either dollar value or physical volume) affects the proper purchasing of production and packaging supplies, the finances needed, transportation facilities for which provision must be made, and influences the selling and advertising effort that must be expended. An accurate sales forecast can, therefore, result in the most efficient utilization (hence the lowest cost) of all the factors involved.

There is a minimum of reference to actual statistical methods employed to achieve accurate sales forecasts. The reader is referred, in the appendix, to some of the standard works on statistical methods and business-cycle theories. The value of this pamphlet lies in its collation and description of the methods in actual use in 107 companies in the United States and Canada. In addition, rather complete case histories of particular methods utilized by 15 representative companies are presented throughout the pamphlet; emphasis is placed on the importance and the necessity of critically examining forecasts based on statistical methods. Judgment, experience, and the test of reasonableness must be applied to every forecast before the forecast becomes an official figure of the company. As the author, in analyzing correlation methods so well puts it, "the more reliable the correlation, the more dangerous its use becomes, since there is a tendency to place blind faith in a method that seems infallible. Even the best correlations are subject to *chance variations*, and one abrupt or severe deviation from normal may be all that is necessary to bring severe losses to a company."

Every company, no matter whether small or large, and whether or not they make official or formal sales forecasts, would be well advised to have this pamphlet read by every employee directly concerned with sales, advertising, production, or financial budgets. No matter how advanced may be the thinking within a particular company with respect to sales forecasting, there is much material in the pamphlet which could probably properly be used in conjunction with the methods now employed by any such company.

Specifically, the pamphlet describes the following methods of sales forecasting, which may include national and territorial sales forecasts by product lines.

NONSTATISTICAL METHODS. (*a*) *The Jury of Executive Opinion.* Under this method, top executives combine their experience and knowledge to arrive at an acceptable figure for budgeting sales. (*b*) *The Sales Force Composite Method.* Estimates are gathered by divisional managers from each of their salesmen and sent to the home office for examination and reconciliation by the top sales-management group in the home office.

STATISTICAL METHODS (Correlation Analysis, Trends and Cycles, and Mathematical Formulas). When statistical methods are employed, the author specifically recommends that such methods be kept simple as possible in order to assure their acceptance by the operating men who are to use the forecasts in the conduct of the company's business. If the factors affecting the sales of a particular company are complex and thus require complex statistical methods, the author recommends that the forecaster "devise means of instilling the necessary understanding and confidence in the ultimate users." Constant review and revision of the forecasts must be made after the period of forecast has begun in order to check the actual performance with the forecasted estimates and to adjust for any change in conditions that may have occurred after the forecasts were made.

MULTIPLE METHOD APPROACH. Under this approach, forecasts made by nonstatistical methods are compared with the forecasts arrived at from statistical analysis. The multiple method is highly recommended since it combines all the experience and knowledge within the company. As the analysis states, "the critical examination of differences between two or more independent forecasts sharpens executive judgment and tends to make the final forecast more reliable."

The only criticism of the pamphlet is the lack of recognition of the use of growth curves for short-term forecasts. Perhaps this method has not yet found wide usage, although its value has long been recognized for long-term forecasts.

It is hoped that the author will find time to expand the material in this pamphlet as more information becomes available, and as the pattern of the postwar economy emerges more clearly. It may well be that prewar statistical correlations, based on a 100 billion dollar maximum economy, may not continue within a 200 billion dollar economy.

Forecasting for Profit: A Technique for Business Management. *Wilson Wright* (Economist, Armstrong Cork Company, Lancaster, Pa.) New York 16: John Wiley & Sons, Inc. (440 Fourth Ave.), 1947. Pp. vii, 173. $2.75. (London, W.C.2, England: Chapman & Hall, Ltd. [37–39 Essex St., Strand], 1948. 16*s.* 6*d.*) *Two reviews follow:*

REVIEW BY WILLIAM A. SPURR

Professor of Business Statistics, Stanford University, Stanford, California

THIS is an essay touching on many topics in business forecasting. The aim is to show the business executive, the business economist and the student how to predict a company's sales, profits, and prices. Toward this end Mr. Wright discusses the need for forecasting, types of forecasts, the factors which determine sales and profits, a method of making a general business forecast, the uses of national income data, the nature of money, the statistical series used by the Armstrong Cork Company, the effect of competition on a company's outlook, business cycles, "economic models," mechanical forecasting devices, types of price forecasting, the determination of turning points in sales and profits, criteria for classifying products, and the final assembly of the sales forecast.

Mr. Wright thus points out the principal problems faced by the forecaster. His discussion appears plausible and persuasive. His viewpoint is that of a practicing business economist. However, I fail to see how one could learn how to make a business forecast from reading this book. There are two fundamental criticisms.

1) The treatment is too superficial. Each problem is merely mentioned in a general way, rather than being carefully analyzed and explained. Furthermore, the topics are heterogeneous and unorganized, and terminology is vague. Definitions are fuzzy, and sometimes inconsistent with ordinary usage. One gets the impression that the book was dictated in a very short time. As further clues to the author's thoroughness, the footnotes, bibliography, and index are rudimentary; the few charts show only hypothetical data, and the very few tables of actual data are brought up only to 1944.

2) There are six essential elements in any good forecasting book; none of which is developed adequately here: (*a*) an explanation of pertinent statistical methods, (*b*) business cycle description and theory, (*c*) a critical examination of leading business indexes, (*d*) analysis of principal noncyclical factors affecting business, (*e*) historical survey of recent economic fluctuations, (*f*) a detailed case study of a specific forecast, showing how the above principles are applied in practice. A business economist should do this last job well, but despite Mr. Wright's many references to his work at the Armstrong Cork Company, he still fails to get across to this reader just how they make a forecast. (The treatment in pages 133–36 is just a start.)

As one illustration of both these criticisms, the statement is made on page 83, that, "... it is helpful to break down or decompose by statistical manipulation an index of business activity into component elements." Very well, then what methods of time series analysis are most appropriate and how may they best be applied in practice? Mr. Wright's only explanation is, "Statististical decomposition involves calculating the various equilibrium values in a series and the mathematical difference of each."

The reader, however, should look at both sides of the matter by contrasting this review with one which appeared in the *American Economic Review* for December 1947 (pp. 974–77). The reviewer there outlines Mr. Wright's thesis at some length and calls it "competent and thorough ... well organized, logical."

To resume: if this treatise is lacking, then what *is* a good book on business forecasting? The prewar books are now rather dated, unfortunately, because of the rapid rate of obsolescence in the data, methods and strategic factors used in forecasting. Three books have been published since the war: (1) Dewey and Dakin's periodicity study, *Cycles: The Science of Prediction* (Henry Holt and Co., Inc., 1947), which is definitely misleading, in my opinion. (2) Wright's *Forecasting for Profit*, which is innocuous but thin. (3) Bratt's *Business Cycles and Forecasting*, Third Edition (Richard D. Irwin, Inc., 1948) which I would recommend as the only adequate treatise of the three.

REVIEW BY VICTOR VON SZELISKI

Vice President, U.S. Economics Corporation
10 Rockefeller Plaza, New York 20, N. Y.

THIS book is recommended to business managers and to professional and academic economists. It is not so much a description of forecasting devices as an inventory and broad description of those parts of the social sciences which are useful to practicing business economists. A more descriptive title would be "Role of Forecasting in Business Management and the Kit of Social Science Tools Needed by the Company Forecaster."

The author correctly regards forecasting as a practical art and craft, in this respect disassociating himself from the school which regards it as a department of mathematics. Forecasting is trying "to understand and describe the specific results of economic and political action taken by human beings.... The tools of the useful practicing economists are statistics of indeterminable accuracy, a knowledge of how people have acted in the past in situations which have been similar in some respects, a disciplined mind, knowledge derived as a result of critical observation, imagination, and common sense." So it is not surprising to find Mr. Wright recommending in the bibliography books by Pareto and Sorokin (but the reviewer does not intend to follow Mr. Wright's bibliography to the extent of reading the book by Major Angas).

The opening pages on the role of forecasting in business today and the position of the forecaster-economist vis-à-vis management are a necessary introduction to any book on practical forecasting.

In a brief paragraph on page 159, the author describes how forecasting has been integrated into the operations of the Armstrong Cork Company through a forecasting committee, how forecasting has been assimilated into the regular operating routine. Perhaps Mr. Wright will expand this tantalizing paragraph at some future time and explain the mechanics of how forecasts are translated into action in purchasing, sales, plant construction, and fiscal management.

The key forecast, in conformity with which an individual company's decisions have to be worked out, is the sales forecast. The author takes up in Chapter 3 the prerequisites necessary for developing a sales forecast, emphasizing the distinction between internal or controllable factors, and external or uncontrollable factors. Analysis of the external factors controlling company sales usually shows that the general business cycle is the dominant factor in company sales.

The first step in business cycle forecasting is making an "estimate of the situation," which is an evaluation of the strains or unbalances currently existing and the direction in which the economy would normally move as a result of these strains. The reviewer's description of the book as a list of needed tools rather than a manual on how to use them is illustrated by the fact that the difficult question of equilibrium and tension is described in less than 4 pages. On the basis of the short description given, the author's method of measuring equilibrium (use of trend lines) appears too mechanical. Equilibrium levels can certainly shift in ways which are not adequately measured by trend lines, and which can be explained statistically only by taking account of a number of factors.

The author's use of expectations is interesting—the more so as reports on the expectations of business executives now appear in *Fortune.*

More hints on diagnosing economic tension are given in Chapter 13. Maladjustments and disorganization within the supply of commodities are looked for; the magnitude of consumers' stocks of durable commodities are to be appraised; their probable influence of prolonged trends are to be evaluated. Important recession and economic activity can be looked for when it is evident that excess productive capacity has been developed in a number of important industries. Residential construction is especially important because of the long, useful life of structures. "A large part of the success experienced in the practice of business forecasting may be attributed to competent analysis of the situation in, and outlook for, the construction industry, with particular emphasis on privately financed residential construction."

It is then up to the forecaster to figure out when and how fast and how far this tension will be released.

As a prerequisite for making useful, short-term forecasts, it is necessary to have a long-term forecast because the latter will condition the interpretation of the short-term occurrences and the prediction of the duration of short-term movements. The interpretation placed upon the expectation of business managers will also depend on the estimation of the long-term disequilibria which have to be corrected.

Increases in labor costs and of material costs per unit of output may be the signal for the release of a tension by a decline in business activity. The author also thinks that stock prices can be profitably analyzed for "measuring the effect of events on the expectations of persons who are highly interested in the trend of business activity and the magnitude of profits." Maybe so. Mr. Wright regards the message of the stock market as especially important when the analysis of tensions indicates that a reversal of the trend of business activity is in the cards.

The author devotes a few pages to economic models. "Being in the vogue, these instruments sometimes have been used to dignify illogical assumption with a mantle of precise arithmetic. Because of this practice, some experienced economists are inclined to lift their eyebrows at the use of these mathematical devices. They can be useful, however, for the purposes of persons who are interested in trying to forecast the development of actual situations in the economy." The reviewer believes that the construction of economic models is absolutely essential for assuring one's self that he has included (or for forcing one's self to include) all major relevant considerations. A model should be thought of as a mosaic picture or a jig-saw puzzle—a matrix into which an economist can fit the thousand-odd clues collected for his work; the individual pieces do not mean very much by themselves but when fitted together they make a picture that is clearer than the mere sum of the parts. In trying to make them fit, the economist learns, and improves his estimate of the situation. At present, however, craftsmanship rather than science must be relied upon in this necessary task because the mathematical methods for model construction still limp badly.

The author also has chapters on industry competition and price forecasting. Raw material prices and selling prices are of immediate interest to industrial concerns, and so also are equity prices, as indications of how easily business can finance new capital expenditures with equity capital—matters which are of extreme importance in those stages of the business cycle where the forecast resolves itself into a question of how much business will spend for capital construction.

The author scarcely discusses at all the opposite problem of how tension and disequilibria are created—how and why the economy moves from a position of equilibrium to one of tension. From the viewpoint of the science of economics, many if not most of the "why's" are accidents; they are political or social in origin.

Chapters 5 through 7 describe measures of national output, national income, the makeup of the money supply, and some of the mechanisms that lead to changes in the money supply. Chapter 7 gives in detail the catalog of statistics in these and related factors, followed by the Armstrong Cork Company as background material from which forecasts are developed.

The author then considers in Chapter 15 how the national business cycle forecast is translated into a forecast of industry and company sales—by correlation techniques, for example. As a preliminary step, the author makes a formal statement of the assumption used in making the various estimates—assumptions drawn from analyses of the general economic pictures, as well as particular assumptions which have to be made about prices of specific commodities with which the company has to be concerned. After preliminary mathematical estimates of company and industrial activity have been set up, it is very important to refine these estimates by considering what deviations from the formula values are likely to obtain. "Guessing" what the residuals are going to be makes a high call on the craftsmanship of the economists. "The most accurate forecasting will be done by those who use judgment as well as mechanical aids in establishing estimated deviations. Personal judgment is required in making a useful estimate of the situation, and judgment again is necessary to establish the expected deviations from computed values when the correlations are used in forecasting for practical purposes."

Deviations from the national pattern may be traced to at least four causes which the professional forecaster must evaluate carefully: (*a*) Secular change in acceptance of the commodity. (*b*) Change in customers' inventories. (*c*) Errors due to inaccuracies in the statistics. (*d*) Inadequacy of the independent variable as a measure of the market.

Among the most interesting pages in the book to a practical forecaster are the summaries or appraisals of economic situations of forecasts prepared for Armstrong Cork Company in 1931, 1932, 1936, 1937, and 1938.

LETTERS ABOUT BOOKS

Readers are invited to submit letters about statistical methodology books for publication in this forum. Concise, informative letters which supplement previously published reviews by pointing out specific strengths, weaknesses, errors, and errata in currently used books are wanted. Criticisms based on actual use of a book as a text are especially desired from statistics instructors. Other letters may consist of suggestions for the writing of books and reviews. Letters which contain adverse criticisms of JOURNAL *reviews will be submitted to the author of the review for any reply he may care to make. Contributors are requested to avoid personalities. The right to decide whether a letter merits publication is reserved. Letters should be sent to the review editor, Oscar K. Buros, Rutgers University, New Brunswick, N. J.*

THEORY OF STATISTICAL INDUCTION

PROFESSOR Kingston has kindly called my attention to the fact that certain of the errors mentioned in my review (this JOURNAL, March 1947) of his book (*A Teoria da Indução Estatistica*, 1945) had, in the meantime, been corrected by the author in *Revista Brasileira da Estatistica* (No. 25, p. 56 and No. 26, p. 326). A rather complete corrigenda reached me after the review was submitted for publication.

I should also like to take this opportunity to clarify the impression that my review has made on some readers. The first paragraph was intended to make it clear that this book was better than some which presumably reputable publishers still add to the large English literature. Though I am not well acquainted with the field, it seems very likely that this is the best and most modern book in any of the Latin languages.

JOHN W. TUKEY, *Assistant Professor of Mathematics, Princeton University, Princeton, N. J.*

EXAMINATION OF INDUSTRIAL MEASUREMENTS

IN THE June 1947 issue of this JOURNAL, two reviews of *Examination of Industrial Measurements* (McGraw-Hill Book Co., 1946) were presented by Professor C. W. Churchman and by Mr. Joseph Manuele and Mr. Roscoe Byers. The writer wishes to comment briefly on these reviews.

Professor Churchman questions our reference to standard "betting odds" in limits recommended by the ASTM and the ASA. Here we are referring to three-sigma limits for indicating assignable variations in control charts and fitted curves. Professor Churchman's statement that "odds differ so widely depending on the product" apparently refers to allowable defectives or allowable quality variations in various products. We hope that he does not infer, for example, that the cement industry is more willing to take chances with its customers than is the textile industry, the steel industry, or any other group of manufacturers.

Professor Churchman's comment on sample sizes from 300 to 600 failed to mention that these were specified for fraction defective only. We agree that this estimate is "arbitrary" along with other such choices made in industrial practice. Actually, our recommendation is based on a statement in L. E. Simon's well-known *Engineer's Manual of Statistical Methods*, and appropriate reference was made in our text.

We agree that the difference between sample estimates and true parameters should be emphasized, and our book follows the ASTM terminology in this respect. However, in industrial work we never know "true" values except as assumed standards. Formulae related to the normal curve (or any other curve) are of no value to the engineer unless he applies them to actual data. In our experience, we have encountered more dangers due to the presence of assignably large variations than those due to sampling error, *assuming that the sample sizes and procedures recommended for control charts are used.*

Professor Churchman's review well illustrates the difference in viewpoint between the statistician and the practicing engineer. Wartime applications have shown that engineers and others of considerably less formal training can grasp and apply the "logical basis" of industrial statistical methods without knowing all of the mathematical derivations. The definition of "logical basis" is a matter of degree, depending upon whether such definition is that of the mathematician, the statistician, the engineer, or the plant foreman.

Mr. Manuele's review states that "neither probability or risk is defined or mentioned" in the discussion of sampling. The three-sigma limit used throughout the book is certainly based on a very definite and strong probability. The engineer who has constructed a set of control charts from his own industrial measurements will thereby become acquainted with the sampling risks incurred, assuming that he has noted the table of odds listed on page 106 of the book.

Although the writer is more interested in applying proven methods than in developing new ones, our description and use of *median-normal limits* and *fraction non-random* are, to the best of our knowledge, new. We shall be interested in references showing where these have been presented elsewhere.

The detailed and helpful nature of these reviews is greatly appreciated by the writer.

JOHN W. DUDLEY, JR., *Chemical Engineer, American Viscose Corporation, 35 South Ninth St., Philadelphia 7, Pa.*

MR. DUDLEY'S comments on my review point out very well the difficulties that inevitably arise in the application of formalized techniques to research and industrial problems. Such techniques are only means to certain objectives of the experimenter or industrial engineer, and may become very bad means, despite their precision, in cases where the conditions of their application are not satisfied. The writer of a practical manual like Mr. Dudley's little volume is faced with the two-fold problem of describing a precise technique for handling measurements, and at the same time advising the reader on the restrictions of application. No one could hope to accomplish this problem completely within the scope of the volume, or even one many times as large. If a reviewer adds a few precautions of his own, this is meant to supplement in some part the author's own comments. Three-sigma limits are *not* always advisable, no fixed sample size is universally applicable, etc. The reader should be made aware as often as possible that the choices he makes result in operating characteristics, which form the real basis for his decisions. And I personally would think this was one of the most "practical" problems in industrial statistics; its solution does not demand a knowledge of mathematical derivations, but rather a knowledge of producer and consumer risks. A knowledge of consumer risks, at least, does take the industrial engineer out of his field, but so much the better.

C. WEST CHURCHMAN, *Associate Professor of Philosophy, Wayne University, Detroit, Michigan.*

STATISTICAL METHODOLOGY INDEX, NO. 11

A QUARTERLY GUIDE TO CURRENT LITERATURE

Edited by
OSCAR KRISEN BUROS
Rutgers University

This bibliographic service presents statistical methodology literature (articles, books, theses, and chapters) published in 1945 to date. Anonymous references are listed first; other references follow in alphabetical sequence by first-named authors. Volume number, issue number (in parentheses), pagination, and date of issue are given for each journal reference. Stars indicate books, theses, and pamphlets; asterisks indicate publications seen by the editor of this INDEX. *For position and address of each author, look up the first entry in which he is the first-listed author. Authors of papers on statistical methodology are invited to send reprints to Oscar K. Buros, Rutgers University, New Brunswick, New Jersey in order to facilitate prompt and accurate listings. Foreign-language papers should be accompanied by English translations of titles. Information is desired about references omitted.*

1377. British Journal of Psychology: Statistical Section. Issued by the British Psychological Society. At least two issues annually; 30s. per volume of about 160 pages; vol. 1, part 1 started October 1947. Edited by CYRIL BURT and GODFREY THOMSON. Published by University of London Press, Ltd., Warwick Square, London E.C.4, England.

1378. Calcutta Statistical Association. *Calcutta Stat Assn B* (1):1–4 Ag '47.*

1379. Calcutta Statistical Association Bulletin. Four issues per year; Rs. 5 or 10s. per year; no. 1 is dated August 1947. P. K. BOSE, Secretary, Calcutta Statistical Association, c/o Department of Statistics, Asutosh Building, Calcutta University, Calcutta 7, India.

1380. ★Conference Papers: First Annual Convention, American Society for Quality Control and Second Midwest Quality Control Conference, Hotel Sherman, Chicago, Illinois, Thursday-Friday, June 5 and 6, 1947. Rochester 8, N. Y.: the Society (c/o Alfred L. Davis, Treas., Rochester Institute of Technology), 1947. Pp. vi, 284. Paper, lithotyped. $3.50. For contents, *see* 1431, 1432, 1438, 1449, 1453, 1454, 1455, 1528, 1562, 1650, 1665, 1691, 1712, 1714, 1718, 1720, 1761, 1767, 1770, 1776, 1781, 1791, 1847, 1848, 1853, 1861.

1381. How Statistical Quality Control Is Being Applied at Timken-Detroit Plant. *Automotive Ind* 97(7):29+ O 1 '47.*

1382. International Actuarial Notation. *Trans Actuarial Soc Am* 48(117):166–76 My '47.*

1383. Mathematical Statistics and Econometrics at the University of North Carolina. *Econometrica* 16(1):125–6 Ja '48.*

1384. New Statistical Method of Predicting Sunspots. *Cur Sci* 16(10):306 O '47.*

1385. [Personnel and Training Problems Created by the Recent Growth of Applied Statistics in the United States.] *Sci* 106 (2756):391–2 O 24 '47.* A summary of 1119.

1386. Quality Control at Cadillac Based on Latest Methods. *Automotive Ind* 97(3): 27+ Ag 1 '47.*

1387. Statistical Methods Simplify Evaluation of Data. *Electronics* 20(11):194+ N '47.*

1388. The Social Use of Sample Surveys. *Planning* (250):1–24 My 24 '46.*

1389. The Teaching of Statistics. *Engineering* 164(4250):38 Jl 11 '47.* Summary of 1753.

1390. The Teaching of Statistics: A Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics. HAROLD HOTELLING, Chairman, WALTER BARTKY (Dean of the Division of the Physical Sciences and Professor of Applied Mathematics, University of Chicago, Chicago, Ill.), W. EDWARDS DEMING (Adviser in Sampling, Bureau of the Budget, Washington, D. C.), MILTON FRIEDMAN, and PAUL HOEL. *Ann Math Stat* 19(1):95–115 Mr '48.*

1391. What Is a Sample? *J Account* 83(6):525 Je '47.*

1392. ★ABBOTT, J. C. (Associate Professor of Mathematics), **and BENAC, T. J.** (Associate Professor of Mathematics). (United States Naval Academy, Annapolis, Md.) **Principles of Counting and Probabil-**

ity. Annapolis, Md.: U. S. Naval Institute, 1947. Pp. iii, 40. Paper. $1.00.* *To be reviewed.*

1393. ABRUZZI, ADAM (Industrial Engineering Consultant, 107 West 107th St., Shanks Village, Orangeburg, N. Y.). **Personnel Records and Statistical Charts.** *Personnel* 24(1):46–54 Jl '47.*

1394. ALBERT, G. E. (Associate Professor of Mathematics, University of Tennessee, Knoxville, Tenn.). **A Note on the Fundamental Identity of Sequential Analysis.** *Ann Math Stat* 18(4):593–6 D '47.*

1395. ALCHIAN, ARMEN A. (Assistant Professor of Economics, University of California, Los Angeles, Calif.). **Part III, Analysis Procedures, pp. 443–509.** In *Records, Analysis, and Test Procedures.* Edited by Walter L. Deemer, Jr. Army Air Forces. Aviation Psychology Program Research Reports, No. 18. Washington, D. C.: Government Printing Office, 1947. Pp. ix, 621. $2.25.*

1396. ANDERSON, J. ANSEL (Chief Chemist, Grain Research Laboratory, Board of Grain Commissioners for Canada, Winnipeg, Manitoba, Canada). **The Preparation of Illustrations and Tables.** *Trans Am Assn Cereal Chem* 3(2):74–104 Ja '45.*

1397. ANDERSON, J. ANSEL. **The Role of Statistics in Technical Papers.** *Trans Am Assn Cereal Chem* 3(2):69–73 Ja '45.*

1398. ANDERSON, R. L. (Associate Professor of Experimental Statistics and Agricultural Economics, North Carolina State College, Raleigh, N. C.). **Use of Variance Components in the Analysis of Hog Prices in Two Markets.** *J Am Stat Assn* 42(240):612–34 D '47.*

1399. ANDERSON, R. L., and ANDERSON, T. W. **Distribution of the Circular Serial Correlation Coefficient for Residuals From a Fitted Fourier Series: Preliminary Report.** Abstract. *Ann Math Stat* 19(1):119 Mr '48.*

1400. ANDERSON, THEODORE W. (Assistant Professor of Mathematical Statistics, Columbia University, New York, N. Y.). and HURWICZ, LEONID (Research Associate Professor, Statistical Laboratory, and Associate Professor of Economics, Iowa State College, Ames, Iowa). **Errors and Shocks in Economic Relationships.** Abstract. *Econometrica* 16(1):36–7 Ja '48.*

1401. ANSCOMBE, F. J. (Lecturer in Mathematics, University of Cambridge, Cambridge, England); GODWIN, H. J. (Lecturer in Mathematics, University College of Swansea, Swansea, Wales); and PLACKETT, R. L. **Methods of Deferred Sentencing in Testing the Fraction Defective of a Continuous Output.** *Sup J Royal Stat Soc* 9(2):198–217 '47.*

1402. ARMITAGE, P. (Statistician, Medical Research Council, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, University of London, London, England). **Some Sequential Tests of Student's Hypothesis.** *Sup J Royal Stat Soc* 9(2):250–63 '47.*

1403. ARMITAGE, P. **A Comparison of Stratified With Unrestricted Random Sampling From a Finite Population.** *Biometrika* 34(3–4):273–80 D '47.*

1404. AROIAN, LEO A. (Assistant Professor of Mathematics, Hunter College, New York 21, N. Y.). **Note on the Cumulants of Fisher's z-Distribution.** *Biometrika* 34(3–4):359–60 D '47.*

1405. AROIAN, L. A., and DARKOW, MARGUERITE (Associate Professor of Mathematics, Hunter College, New York 21, N. Y.). **Fourth Degree Exponential Function.** Abstract. *Ann Math Stat* 18(4):609 D '47.* Also in *B Am Math Soc* 53(11):1128 N '47.*

1406. BAILEY, ARTHUR L. (Statistician, American Mutual Alliance, 60 East 42nd St., New York 17, N. Y.). **A Generalized Theory of Credibility.** *Proc Casualty Actuarial Soc* 32(62):13–20 N 16 '45.*

1407. BAILLIE, DONALD C. (Assistant Professor of Mathematics, University of Toronto, Toronto, Canada). **Actuarial Note: On Testing the Significance of Mortality Ratios by the Use of χ^2.** Discussion by THOMAS N. E. GREVILLE, pp. 541–3; GEORGE C. CAMPBELL, (Assistant Actuary, Metropolitan Life Insurance Co., New York 10, N. Y.), pp. 543–6; CHALMERS L. WEAVER, pp. 546–52; and HAROLD CRAMÉR, p. 552. *Trans Actuarial Soc Am* 47(116, 116S):326–444, 541–61 O '46.*

1408. BAINBRIDGE, J. R. (Commonwealth Research Officer, Mechanical Engineering Department, University of Melbourne, Melbourne, Australia). **The Teaching of Statistics.** Letter. *Engineering* 164(4267):448 N 7 '47.*

1409. BAKER, G. A. (Assistant Professor of Mathematics and Assistant Statistician in the Experiment Station, University of California, Davis, Calif.). **The Effect of Selection Above Definite Lower Limits of Linear Functions of Normally Distributed Correlated Variables on the Means and Variances of Other Linear Functions.** Abstract. *Ann Math Stat* 19(1):118 Mr '48.*

1410. BALDWIN, ALFRED L. (Associate Professor of Psychology, and Research Associate, Samuel S. Fels Research Institute, Antioch College, Yellow Springs, Ohio). **The Study of Individual Personality by Means of the Intraindividual Correlation.** *J Personality* 14(3):151–68 Mr '46.*

1411. BARANKIN, E. W. (Instructor in Mathematics, University of California, Berkeley, Calif.). **Independence of Parameters and Sufficient Statistics.** Abstract. *Ann Math Stat* 19(1):118–9 Mr '48.*

1412. BARNETT, M. K. (Assistant Professor of Chemistry, University of Dayton, Dayton, Ohio). **The Factorial Experiment in Engineering Research.** American Institute of Mining and Metallurgical Engineers, Technical Publication No. 2161. *Metals Technol* 14(4):1–12 Je '47.*

1413. BARTLETT, M. S. (Professor of Mathematical Statistics, University of Manchester, Manchester, England). **The General Canonical Correlation Distribution.** *Ann Math Stat* 18(1):1–17 Mr '47.*

1414. BARTLETT, M. S. **Multivariate Analysis.** *Sup J Royal Stat Soc* 9(2):176–90, discussion 190–7 '47.*

1415. BATES, EDMOND E. (Bates & Boswell, Management Engineers, 1709 West Eighth St., Los Angeles 14, Calif.). **How to Increase Tolerances and Obtain Closer Fits.** *Iron Age* 160(1):58–61 Jl 3 '47.*

1416. BAUMGART, ERNEST L. M. (Laboratoire de Physiologie des Sensations, Collège de France, Paris, France). **The Quantic and Statistical Bases of Visual Excitation.** *J General Physiol* 31(3):269–90 Ja 20 '48.*

1417. BEARD, R. E. (Assistant Fire Manager, Pearl Assurance Co. Ltd., High Holborn, London W.C.1, England). **The Standard Deviation of the Distribution of Sickness.** *J Inst Actuaries Students' Soc* 7(1):23–8 Jl '47.*

1418. BEARD, R. E. **Statistical Problems in Naval Aircraft Provisioning.** *J Inst Actuaries Students' Soc* 6(3):144–8 Ja '47.*

1419. BEERS, HENRY S. (Vice-President, Aetna Life Insurance Co., Hartford, Conn.). **Modified-Interpolation Formulas That Minimize Fourth Differences.** Discussion by WILMER A. JENKINS (Vice-President and Actuary, Teachers Insurance and Annuity Association, New York 18, N. Y.), pp. 184–5; and T. N. E. GREVILLE, pp. 185–6. *Rec Am Inst Actuaries* 34(69, 70):14–20, 184–7 Je, N '45.*

1420. BEERS, HENRY S. **Premium Interpolation.** *Trans Actuarial Soc Am* 48(117):53–75 My '47.*

1421. BELZ, MAURICE H. (Associate Professor of Mathematics, University of Melbourne, Melbourne, Australia). **Note on the Liapounoff Inequality for Absolute Moments.** *Ann Math Stat* 18(4):604–5 D '47.*

1422. BENNETT, W. R. (Research Physicist, Bell Telephone Laboratories, Murray Hill, N. J.). **Distribution of the Sum of Randomly Phased Components.** *Q Appl Math* 5(4):385–93 Ja '48.*

1423. BERKELEY, EDMUND C. (Chief Research Consultant, General Office Administration Department, Prudential Insurance Co., Newark 1, N. J.). **Electronic Machinery for Handling Information, and Its Uses in Insurance.** Discussion by WILLIAM P. BARBER, JR. (Secretary, Connecticut Mutual Life Insurance Co., Hartford, Conn.), pp. 278–80; EDWARD H. WELLS (Assistant Actuary, Mutual Life Insurance Co., New York 5, N. Y.), pp. 280–1; and EDWARD A. RIEDER (Associate Actuary, Mutual Life Assurance Company of Canada, Waterloo, Ontario, Canada), pp. 282–4. *Trans Actuarial Soc Am* 48(117, 118):36–52, 278–88 My, O' 47.*

1424. BHAT, N. R. (Department of Genetics, University of Cambridge, Cambridge, England). **An Improved Genetical Map of Punnett's 'B' Chromosome in the Sweet Pea, *Lathyrus Odoratus* L.** *J Genetics* 48(3):343–58 Ja '48.*

1425. BHATTACHARYYA, A. (Statistical Laboratory, Presidency College, Calcutta, India). **On Some Analogues of the Amount of Information and Their Use in Statistical Estimation.** *Sankhyā* 8(3):201–18 O '47.*

1426. BICKING, CHARLES A. (Quality Control Engineer, Hercules Powder Co., Wilmington 99, Del.). **Quality Control for the Pulp and Paper Industry.** *Paper Trade J* 124(17):63–7 Ap 24 '47.*

1427. BIRNBAUM, Z. W. (Associate Professor of Mathematics, University of Washington, Seattle, Wash.). **On Random Variables With Comparable Peakedness.** *Ann Math Stat* 19(1):76–81 Mr '48.*

1428. BIZLEY, M. T. L. (Technical Assistant, Bacon and Woodrow, Consulting Actuaries, Empire House, St. Martins-le-Grand, London E.C.1, England). **On Waring's B-Formula.** *J Inst Actuaries Students' Soc* 6(3):123–34 Ja '47.* Letter 7(1):53 Jl '47.*

1429. BLAXTER, K. L. (Biochemistry Department, Ministry of Agriculture and Fisheries, Veterinary Laboratory, Weybridge, Surrey, England). **The Evaluation of the Nutritive Value of Animal Feeding-Stuffs: The Application of Statistical Methods to Food Problems.** *Analyst* 73(862):11–5 Ja '48.* Abstract *Chem & Ind* (3):43–4 Ja 18 '47.*

1430. BOCHNER, SOLOMON (Professor of Mathematics, Princeton University, Princeton, N. J.). Stochastic Processes. *Ann Math* 48(4):1014–61 O '47.*

1431. BOLANOVICH, D. J. (Personnel Research Analyst, Personnel Department, Radio Corporation of America, Camden, N. J.). Control Chart Treatment of Personnel Statistics, pp. 111–20. In *Conference Papers. First Annual Convention of American Society for Quality Control. . .* (*see* 1380).

1432. BOLTON, H. R. (Harrison Radiator Division, General Motors Corp., 50 Clyde Ave., Buffalo, N. Y.). Quality Control of Sub-Assemblies, pp. 121–9. In *Conference Papers· First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1433. BOND, W. L. (Member of Technical Staff, Bell Telephone Laboratories, Murray Hill, N. J.). A Simple Procedure for the Making of Alignment Charts. *J Appl Physics* 19(1).83–6 Ja '48.*

1434. BOSE, P. K. (Lecturer in Satistics, Calcutta University, Calcutta, India). On Recursion Formulae, Tables and Bessel Function Populations Associated With the Distribution of Classical D^2-Statistic. *Sankhyā* 8(3):235–48 O '47.*

1435. BOSE, R. C. (Calcutta University, Calcutta, India). On a Resolvable Series of Balanced Incomplete Block Designs. *Sankhyā* 8(3):249–56 O '47.*

1436. BOWEN, G. M. (General Inspector, Feeder Division, Westinghouse Electric Corp., East Pittsburgh, Pa.). Lot Sampling of Screw Machine Parts. *Iron Age* 161(5): 74–6 Ja 29 '48.*

1437. BOWKER, ALBERT H. (Assistant Professor of Mathematical Statistics, Stanford University, Stanford, Calif.). Chap. 2, Tolerance Limits for Normal Distributions, pp. 95–110. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1438. BOYAN, EDWIN A. (Assistant Professor of Business Management, Massachusetts Institute of Technology, Cambridge, Mass.). Is Your Inspection Department Inspecting? pp. 153–9. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1439. BRANDT, A. E. (Statistical Consultant, Office of Technical Director, Naval Ordnance Laboratory, Silver Spring 19, Md.). Principles of Experimental Design Applied to Long-Time Rotations, pp. 306–15. In *Soil Science Society of America Proceedings* 1945, Vol. 10. Morgantown, West Va.: the Society (c/o G. G. Pohlman, Sec.-Treas.), 1946. Pp. viii, 525. $5.00.*

1440. ★BRATT, ELMER CLARK (Professor of Economics, Lehigh University, Bethlehem, Pa.). Business Cycles and Forecasting, Third Edition. Chicago 4, Ill.: Richard D. Irwin, Inc. (332 South Michigan Ave), 1948. Pp. xii, 585. $6.00. *To be reviewed.*

1441. BRATT, ELMER C. Business-Cycle Forecasting. *J Bus* 21(1):1–11 Ja '48.*

1442. ★British Standards Institution. Fraction-Defective Charts for Quality Control. British Standard 1313:1947. London S.W.1: British Standards Institution (28 Victoria St.), 1947. Pp. 40. 6s. Paper.* (New York 17: American Standards Association [70 East 45th St.]. $2.25.) *To be reviewed.*

1443. BRODMAN, ESTELLE (Assistant Librarian), PHEULPIN, FRED J. (Assistant in Charge of Technical Services), (College of Physicians and Surgeons, Columbia University, New York 32, N. Y.); and DEUTSCHBERGER, J. (90 Pinehurst Ave., New York 33, N. Y.). Some Statistical Methods Useful to the Medical Librarian. *B Med Lib Assn* 35(1):7–57 Ja '47.*

1444. BROOKES, B. C. (Modern School, Bedford, Beds, England). The Incorporation of Statistics Into a School Course. *Math Gazette* 31(296):211–8 O '47.*

1445. BROWN, GEORGE W. (Professor of Statistics, Iowa State College, Ames, Iowa). Discriminant Functions. *Ann Math Stat* 18(4):514–28 D '47.*

1446. BROWN, GEORGE W. On Small-Sample Estimation. *Ann Math Stat* 18(4): 582–5 D '47.*

1447. BROWN, GEORGE W., and FLOOD, MERRILL M. (Owner, Merrill Flood and Associates, 20 Nassau St., Princeton, N. J.). Tumbler Mortality. *J Am Stat Assn* 42(240):562–74 D '47.*

1448. BROZEK, JOSEF (Assistant Professor of Physiological Hygiene, University of Minnesota, Minneapolis, Minn.), and ALEXANDER, HOWARD (Professor of Mathematics and Physics, Adrian College, Adrian, Mich.). A Note on Estimation of the Components of Variation in a Two-Way Table. *Am J Psychol* 60(4):629–37 O '47.*

1449. BRUNER, NANCY (Statistician, Western Auto Supply Co., 2107 Grand Ave., Kansas City 8, Mo.). The Fundamentals of Correlation, pp. 259–69. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1450. BURNS, WILBUR L. (Senior Quality Control Analyst and Supervisor of the Statistical Quality Control Section, Harrison Radiator Division, General Motors Corp., Lockport, N. Y.). Quality

Control Proves Itself in Assembly. *Ind Qual Control* 4(4):12–7 Ja '48.*

1451. BUROS, OSCAR KRISEN (Director, Institute of Mental Measurements, and Associate Professor of Education, Rutgers University, New Brunswick, N. J.). **Statistical Methodology Index, No. 9: A Quarterly Guide to Current Literature.** *J Am Stat Assn* 42(239):491–6 S '47.*

1452. BUROS, OSCAR KRISEN. **Statistical Methodology Index, No. 10: A Quarterly Guide to Current Literature.** *J Am Stat Assn* 42(240):668–81 D '47.*

1453. BURR, IRVING W. (Associate Professor of Mathematics, Purdue University, Lafayette, Ind.). **Applications of Control Charts in Industrial Quality Control,** pp. 71–80. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1454. BURR, IRVING W. **General Introduction of Control Chart Principles in Industrial Applications,** pp. 17–27. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1455. BURR, IRVING W., and WESCOTT, MASON E. **Fundamentals of Acceptance Sampling,** pp. 81–95. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1456. BURT, CYRIL (Professor of Psychology, University College, University of London, London W.C.1, England). **A Comparison of Factor Analysis and Analysis of Variance.** *Brit J Psychol, Stat Sect* 1(1):3–26 O '47.*

1457. ★CAMERON, J. M. (Mathematician, Statistical Engineering Laboratory, National Applied Mathematics Laboratories, Bureau of Standards, Washington, D. C.). **Double Sampling Techniques in Acceptance Sampling.** Master's thesis, North Carolina State College (Raleigh, N. C.), 1947.

1458. ★CAMPBELL, FRANCES LOUISE (Professor of Mathematics, George Pepperdine College, Los Angeles 44, Calif.). **A Study of Truncated Bivariate Normal Distributions.** Doctor's thesis, University of Michigan (Ann Arbor, Mich.), 1945.

1459. CARLETON, JOHN (Assistant Actuary, Liberty Mutual Insurance Co., 175 Berkeley St., Boston 16, Mass.). **Non-Random Accident Distributions and the Poisson Series.** *Proc Casualty Actuarial Soc* 32(62):21–6 N 16 '45.*

1460. CARPENTER, H. C. H. (Assistant Actuary, The Marine & General Mutual Life Assurance Society, London E.C.3, England). **Statistical Quality Control.** *J Inst Actuaries Students' Soc* 5(4):207–17 '45.*

1461. CARTER, LAUNOR F. (Assistant Professor of Psychology, University of Rochester, Rochester, N. Y.). **An Experiment on the Design of Tables and Graphs Used for Presenting Numerical Data.** *J Appl Psychol* 31(6).640–50 D '47.*

1462. CARTER, LAUNOR F. Chap. 4, **The Relative Effectiveness of Presenting Numerical Data by the Use of Tables and Graphs,** pp. 65–72. In *Psychological Research on Equipment Design.* Edited by Paul M. Fitts. Army Air Forces Aviation Psychology Program Research Reports, No. 19. Washington, D. C.: Government Printing Office, 1947. Pp. xii, 276. $1.25.*

1463. CATLIN, J. B. (Assistant Technical Director), and STRIEBY, J. G. (Quality Methods Coordinator). (Kimberly-Clark Corp., Neenah, Wis.) **Method for Evaluating and Reporting Product Quality.** *Paper Trade J* 124(15):57–60 Ap 10 '47.*

1464. CATTELL, R. B. (Research Professor of Psychology, University of Illinois, Urbana, Ill.); CATTELL, A. K. S. (Urbana, Ill.); and RHYMER, R. M. (Psychologist, Institute for Juvenile Research, 907 Wolcott, Chicago 12, Ill.). **P-Technique Demonstrated in Determining Psycho-Physiological Source Traits in a Normal Individual.** *Psychometrika* 12(4):267–88 D '47.*

1465. CHAKRABARTI, M. C. (Department of Mathematics and Statistics, Dacca University, Dacca, Pakistan). **On a Special Case of the Distribution Law of the Mean Square Successive Difference.** *B Calcutta Math Soc* 39(1):15–8 Mr '47.*

1466. ★CHAPIN, F. STUART (Professor of Sociology, Chairman of the Department, and Director of the School of Social Work, University of Minnesota, Minneapolis, Minn.). **Experimental Designs in Sociological Research.** New York 16: Harper & Brothers (49 East 33rd St.), 1947. Pp. xi, 206. $3.00. *To be reviewed.*

1467. ★CHEN, WAY MING (Instructor in Mathematics, Brown University, Providence, R. I.). **Power Function of the Analysis of Variance and Covariance of a Normal Bivariate Population.** Doctor's thesis, University of California (Berkeley, Calif.), 1945.

1468. CHUNG, KAI-LAI (Part-time Instructor in Mathematics, Princeton University, Princeton, N. J.). **On a Lemma by Kolmogoroff.** *Ann Math Stat* 19(1):88–91 Mr '48.*

1469. CHUNG, K. L. **On the Maximum Partial Sums of Sequences of Independent Random Variables. Abstract.** *Ann Math*

Stat 18(4):610 D '47. Also in *B Am Math Soc* 53(11):1128–9 N '47.*

1470. ★CHUNG, KAI-LAI. **On the Maximum Partial Sums of Sequences of Independent Random Variables.** Doctor's thesis, Princeton University (Princeton, N. J.), 1947.

1471. CHUNG, KAI-LAI, and ERDÖS, PAUL. **On the Lower Limit of Sums of Independent Random Variables.** *Ann Math* 48(4):1003–13 O '47.*

1472. CLARKE, B. L. (Director of Chemical Control, Merck & Co., Inc., Rahway, N. J.); SMITH, GEORGE F.; YOUDEN, W. J.; SMALLWOOD, HUGH M.; and WILKS, S. S. **Symposium on Statistical Methods in Experimental and Industrial Chemistry.** Discussion by JOHN W. TUKEY, C. WEST CHURCHMAN (Associate Professor of Philosophy, Wayne University, Detroit, Mich.), GRANT WERNIMONT, and JOHN MANDEL (General Physical Scientist, National Bureau of Standards, Washington, D. C.). *Ind & Eng Chem, Anal Ed* 19(12):943–55, discussion 956–60 D '47.* For separate entries of the major papers, *see* 1773, 1776a, 1858, 1874.

1473. CLARKE, R. D. (Prudential Assurance Co., Ltd., Holborn Bars, London E.C.1, England). **An Application of the Poisson Distribution.** *J Inst Actuaries* 72(335):481 Je '46.*

1474. COATES, R. P. (Assistant Actuary), and CODY, DONALD D. (Assistant Actuary). (Equitable Life Assurance Society, New York 1, N. Y.) **Observations on Exposure Sampling Procedures.** Discussion by GEORGE C. CAMPBELL (Assistant Actuary, Metropolitan Life Insurance Co., New York 10, N. Y.), pp. 528–32; CHALMERS L. WEAVER, pp. 532–5; and GORDON D. SHELLARD (Metropolitan Life Insurance Co., New York 10, N. Y.), pp. 535–9. *Trans Actuarial Soc Am* 47(116, 116S):311–25, 528–40 '46.*

1475. COCHRAN, WILLIAM G. (Professor of Experimental Statistics, and Director of Research, Institute of Statistics, North Carolina State College, Raleigh, N. C.). **Recent Developments in Sampling Theory in the United States.** Abstract. *Econometrica* 16(1):71–2 Ja '48.*

1476. COCKFIELD, F. A. (Director of Statistics and Intelligence and Assistant Secretary to the Board of Inland Revenue, Stamps and Taxes Division, Somerset House, London W.C.2, England). **The Distribution of Incomes.** *Economica* 14(56):254–82 N '47.*

1477. COHEN, A. C. (Associate Professor of Mathematics, University of Georgia, Athens, Ga.). **Estimation of Parameters in Truncated Pearson Frequency Distributions.** Abstract. *Ann Math Stat* 18(4):607 D '47.*

1478. COLE, RANDAL H. (Assistant Professor of Mathematics, University of Western Ontario, London, Ontario, Canada). **The Theory of Small Samples: Symposium on Statistics, IV,** pp. 33–6. In *Proceedings of the First Canadian Mathematical Congress, Montreal, 1945.* Toronto, Canada: University of Toronto Press, 1946. Pp. xliv, 367. $3.25.*

1479. COPE, T. F. (Associate Professor of Mathematics, Queens College, Flushing, N. Y.). **What Statistics, If Any, in a Required General Mathematics Course? Report of an Attack on This Problem at Queens College.** Abstract. *Am Math Mo* 55(1):51 Ja '48.*

1480. CORNELL, F. G. (Director, Bureau of Research and Service, College of Education, University of Illinois, Urbana, Ill.). **A Stratified-Random Sample of a Small Finite Population.** *J Am Stat Assn* 42(240):523–32 D '47.*

1481. COULSON, C. A. (Professor of Theoretical Physics, King's College, University of London, London W.C.2, England). **Note on the Random-Walk Problem.** *Proc Cambridge Philos Soc* 43(4):583–6 O '47.*

1482. COWDEN, DUDLEY J. (Professor of Economic Statistics, University of North Carolina, Chapel Hill, N. C.). **Simplified Methods of Fitting Certain Types of Growth Curves.** *J Am Stat Assn* 42(240):585–90 D '47.*

1483. COX, P. R. (Government Actuary's Department, Caxton House East, Tothill, London S.W.1, England). **A Post-War Plan for Population Statistics.** *J Inst Actuaries Students' Soc* 6(1):1–13 Jl '46.*

1484. CRAIG, ALLEN T. (Professor of Mathematics, State University of Iowa, Iowa City, Iowa). **Bilinear Forms in Normally Correlated Variables.** *Ann Math Stat* 18(4):565–73 D '47.*

1485. CRAMÉR, HARALD (Director, Institute of Mathematical Statistics, and Professor of Insurance Mathematics and Mathematical Statistics, University of Stockholm, Stockholm, Sweden). **A Contribution to the Theory of Statistical Estimation.** *Skandinavisk Aktuarietskrift* 29(1–2):85–94 '46.*

1486. CRAWFORD, JAMES R. (Research Economist, Stanford Research Institute, Stanford University, Stanford, Calif.), and HAMMER, PRESTON C. (Associate Professor of Mathematics, Oregon State College, Corvallis, Ore.). **Quality**

Control. Abstract of No. 188. *Metal Prog* 50(1):150+ Jl '46.*

1487. CURTIS, JACK M. (Pharmacologist); UMBERGER, ERNEST J. (Pharmacologist and Chief of the Sex Hormone Section); and KNUDSEN, LILA F. (Division of Pharmacology and Food Division, Food and Drug Administration, Federal Security Agency, Washington, D. C.). **The Interpretation of Estrogenic Assays.** *Endocrinology* 40(4):831–40 Ap '47.*

1488. CURTISS, JOHN H. (Chief, National Applied Mathematical Laboratories, National Bureau of Standards, Washington, D. C.). **Acceptance Sampling by Variables, With Special Reference to the Case in Which Quality is Measured by Average or Dispersion.** *J Res Nat Bur Stand* 39(3):271–90 S '47.* Also available as a separate, Research Paper RP1827, from the Government Printing Office, Washington, D. C. at 15¢ per copy.

1489. DANIELS, H. E. (Lecturer in Mathematics, Clare College, University of Cambridge, England). **Grouping Corrections for High Autocorrelations.** *Sup J Royal Stat Soc* 9(2):245–9 '47.*

1490. DANIELS, H. E., and KENDALL, M. G. **The Significance of Rank Correlations Where Parental Correlation Exists.** *Biometrika* 34(3–4):197–208 D '47.*

1491. ★DANTZIG, GEORGE B. (Mathematical Advisor to the Comptroller, Headquarters, United States Air Forces, Washington, D. C.). **I, Complete Form Neyman-Pearson Fundamental Lemma; II, On the Non-existence of Tests of "Student's" Hypothesis Having Power Functions Independent of Sigma.** Doctor's thesis, University of California (Berkeley, Calif.), 1946.

1492. DAVID, F. N. (Lecturer in Statistics, University College, University of London, London W.C.1, England). **A χ^2 'Smooth' Test for Goodness of Fit.** *Biometrika* 34(3–4):299–310 D '47.*

1493. DAVID, F. N. **A Power Function for Tests of Randomness in a Sequence of Alternatives.** *Biometrika* 34(3–4):335–9 D '47.*

1494. DAVIS, FREDERICK B. (Professor of Psychology, George Peabody College for Teachers; and Director, Test Research Service; Nashville, Tenn.). **The Interpretation of Principal-Axis Factors.** *J Ed Psychol* 38(8):471–81 D '47.*

1495. DAVIS, FREDERICK B. **Appendix B, Technical Note on Test Construction,** pp. 68–79. In his *Utilizing Human Talent: Armed Services Selection and Classification Procedures.* Washington 6, D. C.: American Council on Education (744 Jackson Place), 1947. Pp. ix, 85. Paper. $1.25.*

1496. DELURY, D. B. (Department of Mathematical Statistics, Ontario Research Foundation, Toronto 5, Canada). **On the Estimation of Biological Populations.** *Biometrics* 3(4):145–67 D '47.*

1497. DENSEN, PAUL M. (Chief, Medical Research Statistics Division, Department of Medicine and Surgery, Veterans Administration, Washington, D. C.). **The Development and Use of Statistical Practices in Hospital Work.** *Biometrics* 3(3):109–18 S '47.*

1498. DENSEN, PAUL M. **Statistical Practices in Hospitals.** Condensed from "The Development and Use of Statistical Practices in Hospital Work," published in *Biometrics*, September 1947. *Am Statistician* 1(3):10–22+ D '47.*

1499. DICE, LEE R. (Professor of Zoology and Director of the Laboratory of Vertebrate Zoology, University of Michigan, Ann Arbor, Mich.). **Application of Statistical Methods to the Analysis of Ecologic Association Between Species of Birds.** *Am Midland Nat* 39(1):174–8 Ja '48.*

1500. ★DOUGLASS, RAYMOND D. (Professor of Mathematics), and ADAMS, DOUGLAS P. (Assistant Professor of Graphics). (Massachusetts Institute of Technology, Cambridge, Mass.) **Elements of Nomography.** New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1947. Pp. ix, 209. $3.50.*

1501. DU MAS, FRANK M. (Research Fellow, Department of Psychology, State University of Iowa, Iowa City, Iowa). **On the Interpretation of Personality Profiles.** *J Clin Psychol* 3(1):57–65 Ja '47.*

1502. DUDDING, B. P. (Research Laboratories, General Electric Co. Ltd. of England, Wembly, Middlesex, England). **Control of Quality During Production: An Outline of a New Technique With Some Examples,** pp. 11–12. In *The Times Survey of the British Scientific Instrument Industry.* London E.C.4: Times Publishing Co., Ltd. (Printing House Square), October 1947. Pp. 48. Paper. 1s.*

1503. DUFRENOY, J. (Research Associate, College of Pharmacy, University of California, San Francisco 22, Calif.), and GOYAN, FRANK M. **A Graphical Calculator for Statistical Analysis.** *J Am Pharmaceut Assn, Sci Sect* 36(10):309–14 O '47.*

1504. DUNCAN, ACHESON J. (Associate Professor of Political Economy, Johns Hopkins University, Baltimore, Md.). **Detection of Non-Random Variation When Size of Sample Varies.** *Ind Qual Control* 4(4):9–12 Ja '48.*

1505. EATON, HENRY C. (Field Engineer, American Bosch Corp., 558 Maccabees Bldg., Detroit, Mich.). **Engineered Assembly Tolerance.** *Ind Qual Control* 4(2):16–7 S '47.*

1506. EBEL, ROBERT L. (Assistant Professor of Education, State University of Iowa, Iowa City, Iowa). **The Frequency of Errors in the Classification of Individuals on the Basis of Fallible Test Scores.** *Ed & Psychol Meas* 7(4):725–34 w '47.*

1507. ★EDGE, P. GRANVILLE (Late Lecturer, Division of Medical Statistics, London School of Hygiene and Tropical Medicine, University of London, London, England). **Vital Statistics and Public Health Work in the Tropics Including Supplement on the Genealogy of Vital Statistics.** Foreword by MAJOR GREENWOOD (Professor Emeritus of Epidemiology and Vital Statistics, London School of Hygiene and Tropical Medicine, University of London, London, England). London W.C.2: Baillière, Tindall and Cox (7-8 Henrietta St.), 1947. Pp. xii, 268. (Baltimore 2, Md.: Williams & Wilkins Co. [Mt. Royal & Guilford Avenues]. $5.00.)* *To be reviewed.*

1508. EISENHART, CHURCHILL (Chief, Statistical Engineering Laboratory, National Applied Mathematics Laboratories, National Bureau of Standards, Washington, D. C.). Chap. 4, **Effects of Rounding or Grouping Data,** pp. 185–223. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1509. EISENHART, CHURCHILL. Chap. 8, **Planning and Interpreting Experiments for Comparing Two Standard Deviations,** pp. 267–318. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1510. EISENHART, CHURCHILL. Chap. 14, **Probability That Sample Means Are in Opposite Order to Population Means,** pp. 375–82. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1511. EISENHART, CHURCHILL. Chap. 16, **Inverse Sine Transformation of Proportions,** pp. 395–416. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1512. ★EISENHART, CHURCHILL; HASTAY, MILLARD W. (Research Staff, National Bureau of Economic Research, New York 23, N. Y.); and WALLIS, W. ALLEN; Editors. **Selected Techniques of Statistical Analysis: For Scientific and Industrial Research and Production and Management Engineering.** By the Statistical Research Group, Columbia University, Applied Mathematics Panel, Office of Scientific Research and Development. New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1947. Pp. xiv, 473. $6.00. (London W.C.2: McGraw-Hill Publishing Co., Ltd. [Aldwych House, Aldwych], 1948. 36*s*.)* For contents, *see* 1437, 1508, 1509, 1510, 1511, 1513, 1514, 1515, 1516, 1517, 1543, 1544, 1545, 1607, 1719, 1832, 1833. *To be reviewed.*

1513. EISENHART, CHURCHILL; HASTAY, MILLARD; and WALLIS, W. ALLEN. Chap. 5, **Acceptance Inspection When Lot Quality Is Measured by the Range,** pp. 225–34. Based on work by M. A. GIRSHICK (Douglas Aircraft Co., Santa Monica, Calif.). In *Selected Techniques of Statistical Analysis* (see 1512).

1514. EISENHART, CHURCHILL; HASTAY, MILLARD; and WALLIS, W. ALLEN. Chap. 6, **Variability of Amount of Inspection for Double, Multiple, and Sequential Sampling,** pp. 235–44. Based on work by MYRA LEVINE (Statistician, Socony Vacuum Oil Co., 26 Broadway, New York 4, N. Y.). In *Selected Techniques of Statistical Analysis* (*see* 1512).

1515. EISENHART, CHURCHILL; HASTAY, MILLARD; and WALLIS, W. ALLEN. Chap. 10, **Experiments in Which the Independent Variable Is Known Only in Terms of a Probability Distribution,** pp. 329–38. Based on work by J. WOLFOWITZ (Associate Professor of Mathematical Statistics, Columbia University, New York, N. Y.). In *Selected Techniques of Statistical Analysis* (see 1512).

1516. EISENHART, CHURCHILL; HASTAY, MILLARD; and WALLIS, W. ALLEN. Chap. 12, **Abandoning an Experiment Prior to Completion,** pp. 353–62. Based on work by L. J. SAVAGE (Research Associate, Institute of Radiobiology and Biophysics, and Department of Mathematics, The University of Chicago, Chicago, Ill.), KENNETH J. ARNOLD (Assistant Professor of Mathematics, University of Wisconsin, Madison, Wis.), and M. A. GIRSCHICK (Douglas Aircraft Co., Santa Monica, Calif.). In *Selected Techniques of Statistical Analysis* (*see* 1512).

1517. EISENHART, CHURCHILL, and SOLOMON, HERBERT (Air Intelligence Specialist, Air Intelligence Division, Headquarters, U. S. Air Force, Washington, D. C.). Chap. 15, **Significance of the Largest of a Set of Sample Estimates of Variance,** pp. 383–94. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1518. ELFVING, G. (Professor of Mathematics, University of Helsinki, Helsinki, Finland). **Contributions to the Theory of Integer-Valued Markoff Processes.** *Skandinavisk Aktuarietskrift* 29(3–4):175–205 '46.*

1519. ELFVING, G. **A Simple Method of Deducing Certain Distributions Con-**

nected With Multivariate Sampling. *Skandinavisk Aktuarietskrift* 30(1–2):56–74 '46.*

1520. EPSTEIN, ALBERT (Statistician, International Association of Machinists, Ninth and Mt. Vernon Place, N.W., Washington, D. C.). Statistical Analysis With Hand Punched and Sorted Cards. *Am Statistician* 1(2):6–7 O '47.*

1521. EPSTEIN, BENJAMIN (Coal Research Laboratory, Carnegie Institute of Technology, Pittsburgh, Pa.). The Mathematical Description of Certain Breakage Mechanisms Leading to the Logarithmico-Normal Distribution. *J Franklin Inst* 244 (6):471–7 D '47.*

1522. EPSTEIN, BENJAMIN. The Role of the Mellin Transform in the Study of the Distribution of the Product and Quotient of Independent Random Variables. Abstract. *B Am Math Soc* 54(1):75 Ja '48.*

1523. ERDÓS, P. (Research Professor of Mathematics, Syracuse University, Syracuse, N. Y.), and KAC, M. On the Number of Positive Sums of Independent Random Variables. *B Am Math Soc* 53(10):1011–20 O '47.*

1524. EVANS, A. W. (Actuary and Manager, London Life Association, Ltd., 81 King William St., London E.C.4, England). Further Remarks on the Relationship Between the Values of Life Annuities at Different Rates of Interest, Including a Description of a Method of First-Difference Interpolation and a Reference to Annuities-Certain. *J Inst Actuaries* 72(335):447–54 Je '46.*

1525. FAY, ERNEST L. (Quality Control Advisor, Deere & Co., Moline, Iowa). Common-Sense Quality Control. *Am Mach* 91(20–21):76–8, 110–3 S 25, O 9 '47.*

1526. FAY, ERNEST L. Purchased-Part Quality Controlled. *Am Mach* 91(26): 98–100 D 18 '47.*

1527. FERRELL, ENOCH B. (Bell Telephone Laboratories, Inc., 463 West St., New York 14, N. Y.). Statistical Methods in the Development of Apparatus Life Quality. *Trans Am Inst Elec Eng* 64:700–2, discussion 998–9 '46.*

1528. FERRIS, DOUGLAS W. (Quality Control Engineer, Nickey Brothers, Inc., Memphis, Tenn.). Application of Quality Control in Woodworking, pp. 213–20. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1529. FESTINGER, LEON (Research Center for Group Dynamics, Massachusetts Institute of Technology, Cambridge, Mass.). The Treatment of Qualitative Data by "Scale Analysis." *Psychol B* 44(2):149–61 Mr '47.*

1530. FINDLEY, WARREN G. (Chief, Evaluation Unit, Educational Advisory Staff, Air University, Maxwell Field, Ala.). A Statistical Index of Participation in Discussion. *J Ed Psychol* 39(1):47–51 Ja '48.*

1531. FINNEY, D. J. (Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford, Oxford, England). The Estimation From Individual Records of the Relationship Between Dose and Quantal Response. *Biometrika* 34(3–4): 320–34 D '47.*

1532. FINNEY, D. J. The Inevitability of Statistics: The Application of Statistical Methods to Food Problems. *Analyst* 73 (862):2–6 Ja '48.* Abstract *Chem & Ind* (3):43–4 Ja 18 '47.*

1533. FINNEY, D. J. The Significance of Associations in a Square Point Lattice. *Sup J Royal Stat Soc* 9(1):99–103 '47.*

1534. FINNEY, D. J. Statistical Aspects of Microbiological Assays. Abstract. *Biochem J* 41(1):v–vii '47.*

1535. FISHBACK, PHILIP G. (Quality Control Engineer, Graflex, Inc., Rochester 4, N. Y.). Planned Simplified Inspection. *Tool Eng* 17(4):30–4 N '46.*

1536. FISHBACK, PHILIP G. Screening vs. Sampling in Inspection. *Tool Eng* 18(5):18–23 Je '47.*

1537. FISHER, J. C. (Research Associate), and HOLLOMAN, J. H. (Research Metallurgist). (General Electric Co., Schenectady, N. Y.) A Statistical Theory of Fracture. American Institute of Mining and Metallurgical Engineers, Technical Publication No. 2218. *Metals Technol* 14(5):1–16 Ag '47.*

1538. FISHER, R. A. (Arthur Balfour Professor of Genetics, University of Cambridge, Cambridge, England). A System of Scoring Linkage Data, With Special Reference to the Pied Factors in Mice. *Am Nat* 80(794):568–78 S-O '46.*

1539. ★FISHER, R. A. The Theory of Linkage in Polysomic Inheritance. *Philos Trans Royal Soc London, Series B, Biol Sci* 233(594):55–87 Je 19 '47.* (London N.W.1: Cambridge University Press [Bentley House]. 5*s.* 6*d.*)

1540. FIZZELL, JAMES A. (Supervisor of the Engineering Department, Illinois Testing Laboratories, Inc., 420 N. LaSalle St., Chicago 10, Ill.). Let's Try Analysis of Variance. *Ind Qual Control* 4(4):22–5 Ja '48.*

1541. FORSTER, C. I. K. (Statistical Adviser to the Director of Air Equipment, Admiralty, London S.W.1, England). An Elementary Introduction to the Testing of Statistical Hypotheses. *J Inst Actuaries Students' Soc* 7(2):81–97 O '47.*

1542. FOURIE, J. C. **Sampling, Graphical Representation and Block Valuation.** Discussion by S. E. WILLOWS-MUNRO and B. J. KLOPPERS. *J Chem Metall & Mining Soc South Africa* 48(5):129–43, discussion 143–8 N '47.*

1543. FRIEDMAN, MILTON (Associate Professor of Economics, University of Chicago, Chicago, Ill.). **Chap. 9, Utilization of Limited Experimental Facilities When the Cost of Each Measurement Depends on Its Magnitude,** pp. 319–28. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1544. FRIEDMAN, MILTON. **Chap. 11, Planning an Experiment for Estimating the Mean and Standard Deviation of a Normal Distribution From Observations on the Cumulative Distribution,** pp. 339–52. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1545. FRIEDMAN, MILTON, and SAVAGE, L. J. **Chap. 13, Planning Experiments Seeking Maxima,** pp. 363–72. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1546. GAGE, N. L. (Associate Director), and REMMERS, H. H. (Professor of Education and Psychology and Director). (Division of Educational Reference, Purdue University, Lafayette, Ind.) **Opinion Polling With Mark-Sensed Punch Cards.** *J Appl Psychol* 32(1):88–91 F '48.*

1547. GEARY, R. C. (Director of Statistics, Statistics Branch, Department of Industry and Commerce, Dublin, Eire). **Sampling Aspects of the Problem of Relationship From the Error-in-Variable Approach.** Abstract. *Econometrica* 16(1):37–8 Ja '48.*

1548. GEARY, R. C. **Testing for Normality.** *Biometrika* 34(3–4):209–42 D '47.*

1549. GEIRINGER, HILDA (Professor of Mathematics, Wheaton College, Norton, Mass.). **Contribution to the Heredity Theory of Multivalents.** *J Math & Physics* 26 (4):246–78 Ja '48.*

1550. GENGERELLI, J. A. (Associate Professor of Psychology, University of California, Los Angeles, Calif.). **A Method of Binomial Analysis.** Abstract. *Am Psychol* 2 (10):406 O '47.*

1551. GEORGESCU-ROEGEN, NICULAE (Professor, Institute of Statistics, Actuarial Studies, and Calculus, and Director, Ministry of National Economy, Bucharest, Rumania). **Further Contributions to the Scatter Analysis.** Abstract. *Econometrica* 16(1):40–3 Ja '48.*

1552. GERMOND, HALLETT H. (Director of Research, S. W. Marshall & Co., Consulting Engineers, Washington, D. C.). **Circular Probable Error of an Elliptical Gaussian Distribution.** Abstract. *Ann Math Stat* 19(1):121 Mr '48.*

1553. GHOSH, BIRENDRANATH (Presidency College, Calcutta, India). **Bias Introduced by Changing the System of Stratification.** *Calcutta Stat Assn B* (1):43–5 Ag '47.*

1554. GHOSH, BIRENDRANATH. **Crop Estimation in India: A Brief Review of the Sampling Methods.** *Calcutta Stat Assn B* (1):5–12 Ag '47.*

1555. GHOSH, BIRENDRANATH. **Double Sampling With Many Auxiliary Variates.** *Calcutta Stat Assn B* (2):91–3 N '47.*

1556. GODDESS, EUGENE (Head of Factory Engineering, Cathode Ray Tube Division, North American Philips Co., Inc., Dobbs Ferry, N. Y.). **Charts Help Diagnose Quality Progress.** *Am Mach* 91:96–100 Je 5 '47.*

1557. GOFFMAN, CASPER (Associate Professor of Mathematics, University of Oklahoma, Norman, Okla.). **Statistics as an Aid to Engineering Judgment in the Manufacture of Lightning-Arrester Blocks.** *Trans Am Inst Elec Eng* 64:607–9, 966–8 '46.*

1558. GOOD, RAY W. (Assistant Chief Inspector, Robertshaw-Fulton Controls Co., Youngwood, Pa.). **Quality Control in Relation to Time-Study.** *Ind Qual Control* 4(2):13–6 S '47.*

1559. GOODE, HENRY P. (Associate Professor of Mechanical Engineering, Stanford University, Stanford, Calif.). **How, Why, When, Where, and How Many? A New Approach to Accident Frequency Rates.** *Safety Eng* 92(1):25–6+ Jl '46 *

1560. GOODENOUGH, FLORENCE (Research Professor, Institute of Child Welfare, University of Minnesota, Minneapolis, Minn.). **Note on an Unnecessary Source of Confusion in Statistical Terminology.** *J Ed Psychol* 38(7):443–5 N '47.*

1561. GORE, W. L. (Research Group Leader, Plastics Department, E. I. Du Pont De Nemours and Co., Inc., Arlington, N. J.). **Quality Control in the Chemical Industry: IV, Statistical Methods in Plastics Research and Development.** *Ind Qual Control* 4(2): 5–8 S '47.*

1562. GORE, W. L. **Statistical Techniques for Research and Development,** pp. 145–51. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1563. GOYAN, FRANK M. (Assistant Professor of Chemistry, College of Pharmacy, University of California, San Francisco 22, Calif.), and DUFRENOY, J. **A Graphical Calculator for Bioassays.** *J Am*

Pharmaceut Assn, Sci Sect 36(10):305–8 O '47.*

1564. GREENFIELD, E. W. (Engineer-in-Charge, Electrical Laboratory, Anaconda Wire and Cable Co., Hastings-on-Hudson, N. Y). Refresher on Statistical Analysis Applied to Two A.S.T.M. Oil Dielectric Strength Test Procedures. *ASTM B* (149):89–94 D '47.*

1565. GREVILLE, T. N. E. (Principal Actuarial Mathematician, National Office of Vital Statistics, U. S. Public Health Service, Washington, D. C.). Actuarial Note: Some Extensions of Mr. Beer's Method of Interpolation. Discussion by HENRY S. BEERS, pp. 188–91. *Rec Am Inst Actuaries* 34(69, 70):21–34, 188–93 Je, N '45.*

1566. GREVILLE, T. N. E. Remark on the Note "A Generalization of Waring's Formula." *Ann Math Stat* 18(4):605–6 D '47.*

1567. GREVILLE, T. N. E., pp. 35–9; WELLS, E. H. (Assistant Actuary, Mutual Life Insurance Co., New York 5, N. Y.); pp. 39–43; CAMPBELL, GEORGE C. (Assistant Actuary, Metropolitan Life Insurance Co., New York 10, N. Y.), pp. 43–6; JENKINS, WILMER A., pp. 46–50; and BERGSTRESSER, R. L. (Assistant Actuary, U. S. Life Insurance Co., New York 7, N. Y.), pp. 50–52. Discussion of "Six-Term Formulas for Routine Actuarial Interpolation" by Henry S. Beers. Reply by HENRY S. BEERS, pp. 52–61. *Rec Am Inst Actuaries* 34(69):35–61 Je '45.*

1568. GROSS, MORTIMER D. (Formerly Research Chemist, University of Akron and the United States Rubber Co.; 1361 East 57th St., Chicago 37, Ill.). Quality Control in the Chemical Industry: V, Quality Control in the Synthetic Rubber Industry. *Ind Qual Control* 4(3):11–6 N '47.*

1569. GUILFORD, J. P. (Professor of Psychology, University of Southern California, Los Angeles, Calif.). Approaches to the Univocal Factor Scores. Abstract. *Am Psychol* 2(10):405 O '47.*

1570. GUMBEL, E. J. (Associate Professor of Mathematics, Brooklyn College, Brooklyn, N. Y.). Probability of Death and Expectation of Life. *Human Biol* 18(4):238–40 D '46.*

1571. GURLAND, J. (Research Assistant and Associate, Department of Mathematics, University of California, Berkeley, Calif.). An Inversion Formula for the Distribution of a Ratio of Random Variables. Abstract. *Ann Math Stat* 19(1):118 Mr '48.*

1572. HAAVELMO, TRYGVE (Chief of The National Economic Budget Bureau, Department of Commerce, Oslo, Norway). Family Expenditures and the Marginal Propensity to Consume. *Econometrica* 15(4): 334–41 O '47.*

1573. HAEMER, KENNETH W. (Statistician, American Telephone and Telegraph Co., 195 Broadway, New York 7, N. Y.). The Perils of Perspective. *Am Statistician* 1(3):19 D '47.*

1574. HAGOOD, MARGARET JARMAN. (Principal Statistician, Division of Farm Population and Rural Life, Bureau of Agricultural Economics, U. S. Department of Agriculture, Washington, D. C.). Recent Contributions of Statistics to Research Methodology in Sociology. *Social Forces* 26 (1):36–40 O '47.*

1575. ★HALD, A. (Assistant Professor, University of Copenhagen, Copenhagen, Denmark). The Decomposition of a Series of Observations Composed of a Trend, a Periodic Movement, and a Stochastic Variable. Copenhagen, Denmark: G.E.C. Gads Forlag, 1948. Pp. 134. Paper. *To be reviewed.*

1576. HALDANE, J. B. S. (Weldon Professor of Biometry, University College, University of London, London W.C.1, England). The Dysgenic Effect of Induced Recessive Mutations. *Ann Eug* 14(1):35–43 O '47.*

1577. HALDANE, J. B. S. The Theory of a Cline. *J Genetics* 48(3): 277–84 Ja '48.*

1578. HALDANE, J. B. S., and LEA, D. E. (Lately of The Strangeways Research Laboratory, Mills Road, Cambridge, England). A Mathematical Theory of Chromological Rearrangements. *J. Genctics* 48(1):1–10 Ap '47.*

1579. HALDANE, J. B. S., and SMITH, C. A. B. (Assistant Lecturer in Eugenics, University College, University of London, London W.C.1, England). A New Estimate of the Linkage Between the Genes for Colour-Blindness and Haemophilia in Man. *Ann Eug* 14(1): 10–31 O '47.*

1580. HALE, W. T. (Statistical Section, Research and Development Department, United Steel Companies Ltd., Sheffield 10, England). A Statistical Sampling Plan for Refractory Products With Special Reference to Silica Bricks. *Brit Ceramic Soc Trans* 46(4):147–60 Je '47.*

1581. HANSEN, MORRIS H. (Statistical Assistant to the Director, Bureau of the Census, Washington, D. C.). Sampling of Human Populations. Abstract. *Econometrica* 16(1): 74–5 Ja '48.*

1582. HARRIES, HU (Department of Railways and Telephones, Edmonton, Alberta, Canada). The Development and Use of Production Functions for Firms in Agriculture. *Sci Ag* 27(10):487–93, discussion 493–5 O '47.*

1583. HARRIS, T. E. (Statistical Department, Douglas Aircraft Co., Santa Monica, Calif.). Some Further Results on the Bernoulli Process. Abstract. *Ann Math Stat* 19(1):116 Mr '48.*

1584. ★HARRIS, THEODORE E. Some Theorems on the Bernoullian Multiplicative Process. Doctor's thesis, Princeton University (Princeton, N. J.), 1947.*

1585. HARRIS, T. E. Some Theorems on the Bernoullian Multiplicative Process. Abstract. *Ann Math Stat* 18(4):611 D '47.*

1586. ★HARSHBARGAR, BOYD (Professor of Statistics, and Statistician, Agricultural Experiment Station, Virginia Polytechnic Institute, Blacksburg, Va.). Rectangular Lattices. Virginia Agricultural Experiment Station, Memoir 1. Blacksburg, Va.: the Station, October 1947. Pp. iii, 26. Paper.*

1587. HARTLEY, H. O. (Lecturer in Statistics), and KHAMIS, S. H. (Palestine Government Research Scholar). (University College, University of London, London W.C.1, England.) A Numerical Solution of the Problem of Moments. *Biometrika* 34 (3–4):340–51 D '47.*

1588. HARTMAN, JOHN (Professor in Vegetable Crops, Cornell University, Ithaca, N. Y.), and STAIR, E. C. (Associate Professor of Horticulture, Purdue University, Lafayette, Ind.). Correlation of Means and Standard Deviations in Tomato Field Experiments. pp. 337–40. In *Proceedings of the American Society for Horticultural Science*, Vol. 48. Geneva, N. Y.: W. F. Humphrey Press, December 1946. Pp. viii, 522. $6.00.*

1589. ★HATKE, SISTER M. AGNES (Dean of Students, St. Francis College, Fort Wayne 8, Ind.). A Certain Cumulative Probability Function. Doctor's thesis, Purdue University (Lafayette, Ind.), 1947.

1590. HAYCOCKS, H. W. (Secretary, Actuarial Tuition Service, 6 Staple Inn, Holborn, London W.C.1, England). Probability. *J Inst Actuaries Students' Soc* 6(4): 189–95 Ap '47.*

1591. HEIDE, J. D. (General Laboratories, U. S. Rubber Co., Passaic, N. J.). Statistical Control in the Physical Testing of Rubber. *India Rubber World* 114(5):653–6+ Ag '46.*

1592. ★HILL, A. BRADFORD (Professor of Medical Statistics and Director of the Department, London School of Hygiene and Tropical Medicine, University of London, London, England). Principles of Medical Statistics, Fourth Edition. London W.C.2: Lancet Ltd. (7 Adams St., Adelphi), 1948. Pp. xi, 252. 10s. 6d.* *To be reviewed.*

1593. HILL, A. BRADFORD. Statistics in Medicine. *Trans Manchester Stat Soc* 1946–47:1–15 '47.*

1594. HOBSON, L. S. (Assistant Works Superintendent); INGLIS, R. S. (Superintendent of Quality Control and Inspection); and McCANTS, R. P. (Welding Engineer). (General Electric Co., 6901 Elmwood Ave., Philadelphia 42, Pa.) Application of Quality Control to Resistance Welding. *Trans Am Inst Elec Eng* 64: 573–5, discussion 951–2 '46.*

1595. HOEFFDING, WASSILY (Research Associate, Department of Mathematical Statistics, University of North Carolina, Chapel Hill, N. C.). The Power of Certain Non-Parametric Tests of Independence. Abstract. *Ann Math Stat* 18(4): 607 D '47.*

1596. HÖFFDING [HOEFFDING], WASSILY. On the Distribution of the Rank Correlation Coefficient τ When the Variates Are Not Independent. *Biometrika* 34(3–4): 183–96 D '47.*

1597. HOEL, PAUL G. (Associate Professor of Mathematics, University of California, Los Angeles, Calif.). Discriminating Between Bionomial Distributions. *Ann Math Stat* 18(4):556–64 D '47.*

1598. HOEL, PAUL G. On the Choice of Forecasting Formulas. *J Am Stat Assn* 42 (240):605–11 D '47.*

1599. HOEL, PAUL G. On the Selection of Forecasting Formulas. Abstract. *Ann Math Stat* 19(1):117 Mr '48.*

1600. HOEL, PAUL G. On the Uniqueness of Similar Regions. *Ann Math Stat* 19(1):66–71 Mr '48.*

1601. HOLIDAY, E. R. (The London Hospital, Whitechapel Road, London E.3, England), and IRWIN, J. O.; assisted by A. E. GILLAM (Department of Chemistry, Victoria University, Manchester, England), R. A. MORTON (Johnston Professor of Biochemistry, The University of Liverpool, Liverpool, England), T. MOORE (Dunn Nutritional Laboratory, Milton Road, Cambridge, England), and Y. L. WANG (National Central University Medical College, Chengtu, Szechuan, China). The Assay of Crystalline Preparations of Aneurine by Means of the Ultra-Violet Absorption Spectrum. *Q J Pharm & Pharmacol* 19(2):155–72 Ap-Je '46.*

1602. HOPKINS, J. W. (Biometrician, Division of Applied Biology), and WEATHERBURN, MURIEL W. (Chemist, Division of Chemistry). (National Research Laboratories, Ottawa, Canada.) Precision of Laboratory Measurements of Breaking

Strength of Textiles. *Can J Res, Sect F* 25 (4):264–72 Jl '47.*

1603. HORNSETH, RICHARD A. (Instructor in Sociology, University of Wisconsin, Madison, Wis.). **A Note on "The Measurement of Ecological Segregation" by Julius Jahn, Calvin F. Schmid, and Clarence Schrag.** *Am Sociol R* 12(5):603–4 O '47.*

1604. HORST, A. P. (Professor of Psychology, University of Washington, Seattle, Wash.). **Appendix D, The Relationship Between the Validity of a Single Test and Its Contribution to the Predictive Efficiency of a Test Battery,** pp. 603–8. In *Records, Analysis, and Test Procedures.* Edited by Walter L. Deemer, Jr. Army Air Forces, Aviation Psychology Program Research Reports, No. 18. Washington, D. C.: Government Printing Office, 1947. Pp. ix, 621. $2.25.*

1605. HORTON, H. BURKE (Senior Transport Analyst, Interstate Commerce Commission, Washington, D. C.). **A Method for Obtaining Random Numbers.** *Ann Math Stat* 19(1):81–5 Mr '48.*

1606. HOTELLING, HAROLD (Professor of Mathematical Statistics and Head of the Department; Associate Director, Institute of Statistics; University of North Carolina, Chapel Hill, N. C.). **Effects of Non-Normality at High Significance Levels.** Abstract. *Ann Math Stat* 18(4):608–9 D '47.* Also in *B Am Math Soc* 53(11):1129 N '47.*

1607. HOTELLING, HAROLD. **Some New Methods for Distributions of Quadratic Forms.** Abstract. *Ann Math Stat* 19(1):119 Mr '48.*

1608. HOTELLING, HAROLD. **The Training of Statisticians.** A summary of a report by the Committee on the Teaching of Statistics of the Institute of Mathematical Statistics. *Am Statistician* 1(3):8–9 D '47.*

1609. HOTELLING, HAROLD. **Chap. 3, Multivariate Quality Control, Illustrated by the Air Testing of Sample Bombsights,** pp. 111–84. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1610. HOUSEMAN, EARL E. (Statistical Consultant, Bureau of Agricultural Economics, Washington, D. C.). **The Sample Design for a National Farm Survey by the Bureau of Agricultural Economics.** *J Farm Econ* 29(1):241–5 F '47.*

1611. HSU, E. H. (Instructor in Psychology, Catholic University of America, Washington, D. C.). **A New Method for Determining Reliability Coefficient: The Consistency Index.** *J General Psychol* 37(2):187–91 O '47.*

1612. HSU, P. L. (Associate Professor of Mathematical Statistics, University of North Carolina, Chapel Hill, N. C.). **A General Weak Limit Theorem for Independent Distributions.** Abstract. *Ann Math Stat* 18(4):609–10 D '47.* Also in *B Am Math Soc* 53(11):1129 N '47.*

1613. HULL, FRED H. (Agronomist, Florida Agricultural Experimental Station, Gainesville, Fla.). **Theoretical Regression of F_N on Homozygous Parents: With Additive or Complementary Gene Action.** Abstract. *Genetics* 32(1):91 Ja '47.*

1614. HULL, FREDERICK C. (Research Metallurgist, Westinghouse Research Laboratories, East Pittsburgh, Pa.). **A New Method for Making Rapid and Accurate Estimates of Grain Size.** American Institute of Mining and Metallurgical Engineers, Technical Publication No. 2160. *Metals Technol* 14(4):1–13 Je '47.*

1615. HUTCHINSON, G. E. (Professor of Zoology, and Scientific Consultant in the Bingham Oceanographic Laboratory, Yale University, New Haven, Conn.). **A Note on the Theory of Competition Between Two Social Species.** *Ecol* 28(3):319–21 Jl '47.*

1616. ★Inter-American Statistical Institute. **Bibliography of Selected Statistical Sources of the American Nations: A Guide to the Principal Statistical Materials of the 22 American Nations, Including Data, Analyses, Methodology, and Laws and Organization of Statistical Agencies.** Washington 25, D. C.: Inter American Statistical Institute, Halbert L. Dunn, Sec. Gen. (c/o National Office of Vital Statistics), 1947. Pp. xvi, 689.* *To be reviewed.*

1617. IRWIN, J. O. (Lecturer in Medical Statistics, London School of Hygiene and Tropical Medicine, University of London, London, England); assisted by NANCY GOODMAN. **The Statistical Treatments of Measurements of the Carcinogenic Properties of Tars (Part I) and Mineral Oils (Part II).** *J Hyg* 44(5):362–420 My '46.*

1618. IYER, P. V. KRISHNA (Lectureship in the Design and Analysis of Scientific Experiment, University of Oxford, Oxford, England). **Random Association of Points on a Lattice.** Letter. *Nature* 160(4073):714 N 22 '47.*

1619. JACKSON, ROBERT W. (President, Aldens' Chicago Mail Order Co., Inc., Chicago, Ill.). **Quality Control at Aldens,** pp. 11–5. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1620. JACOBSON, PAUL H. (Vital Statistics Analyst, Statistical Bureau, Metropolitan Life Insurance Co., New York 10, N. Y.). **The Relative Power of Three Statistics for Small Sample Destructive Tests.** *J Am Stat Assn* 42(240):575–84 D '47.*

1621. JAMES, HUBERT M. (Professor of Physics, Purdue University, Lafayette, Ind.). Statistical Properties of Networks of Flexible Chains. *J Chem Physics* 15(9) :651–68 S '47.*

1622. JARCHO, SAUL (Adjunct Physician, Mount Sinai Hospital, New York 29, N. Y.). Equal-Area Projections and the Azimuthal Equidistant Projection in Maps of Disease. *Am J Pub Health* 35(10) :1005–13 O '45.*

1623. JARRETT, R. F. (Instructor in Psychology, University of California, Berkeley, Calif.). Statistical Criteria of the Effectiveness of Selective Procedures. Abstract. *Am Psychol* 2(10) :404 O '47.*

1624. JOACHIM, GERTRUD STANTON. (Research Fellow, Division of Agronomy and Plant Genetics, University of Minnesota, St. Paul, Minn.). The Product Method of Calculating Linkage From F_2 Data Involving Semisterility; and Its Application to a Barley Translocation. *Genetics* 32(6) :580–91 N '47.*

1625. ★JOHNSEN, MADELINE M. (Instructor in Mathematics, Purdue University, Lafayette, Ind.). Approximate Evaluation of Double Probability Integrals. Doctor's thesis, Stanford University (Stanford, Calif.), 1946.

1626. JOHNSON, PALMER O. (Professor of Education, University of Minnesota, Minneapolis, Minn.), and HOYT, CYRIL (Research Associate in Education, University of Chicago, Chicago, Ill.). On Determining Three Dimensional Regions of Significance. *J Exp Ed* 15(4) :342–53 Je '47.*

1627. JONES, DALE (Senior Industrial Engineer, Johns-Manville Corporation, Manville, N. J.). Simple Way to Figure Machine Downtime. *Factory Mgmt* 104(10): 118–21 O '46.*

1628. JOSEPH, A. W. (Assistant Actuary, Wesleyan and General Assurance Society, Steelhouse Lane, Birmingham 4, England). Actuarial Examinations. *J Inst Actuaries Students' Soc* 6(3) :155 Ja '47.* Letter regarding 1686.

1629. JOSEPH, A. W. A Problem in Derangements. *J Inst Actuaries Students' Soc* 6(1) :14–22 Jl '46.* *See also* 1685.

1630. JOSEPH, A. W. The Valuation of the Whole-Life Assurances by the Use of Moments. J *Inst Actuaries* 72(335) :498–515 Je '46.*

1631. KAC, M. (Associate Professor of Mathematics, Cornell University, Ithaca, N. Y.). On the Average of a Certain Wiener Functional and a Related Theorem in Calculus of Probability. *Trans Am Math Soc* 59(43) :401–14 My '46.*

1632. KAC, MARK. On the Distribution of Certain Wiener Functionals: Preliminary Report. Abstract. *B Am Math Soc* 54(1): 64 Ja '48.*

1633. KAC, M. On the Notion of Recurrence in Discrete Stochastic Processes. *B Am Math Soc* 53(10) :1002–10 O '47.*

1634. KARPINOS, BERNARD D. (Major, Medical Department, and Chief Medical Statistician, Surgeon General's Office, Washington, D. C.). Use of Life Table Death Rates for Comparative Mortality. *Human Biol* 18(2) :127–31 My '46.*

1634a. ★KATZ, LEO (Assistant Professor of Mathematics, Michigan State College, East Lansing, Mich.). Characteristics of Frequency Functions Defined by First Order Difference Equations. Doctor's thesis, University of Michigan (Ann Arbor, Mich.), 1946.

1635. KATZ, LEO. Frequency Functions Defined by the Pearson Difference Equation. Abstract. *Ann Math Stat* 19(1) :120 Mr '48.*

1636. KAUFFMAN, L. S. (Quality Control Director, American Stove Co., 1641 S. Kingshighway Blvd., St. Louis 10, Mo.). Statistical Methods Applied to the Inspection of Porcelain Enamel. *Ceramic Ind* 49 (2) :61–2 Ag '47.* Reprinted from *Ind Qual Control* N '46.

1637. KEEPING, E. S. (Assistant Professor of Mathematics, University of Alberta, Edmonton, Alberta, Canada). Estimation and Confidence: Symposium on Statistics, V, pp. 36–42. In *Proceedings of the First Canadian Mathematical Congress*, Montreal, 1945. Toronto, Canada: University of Toronto Press, 1946. Pp. xliv, 367. $3.25.*

1638. ★KELLEY, TRUMAN LEE (Professor of Education, Harvard University, Cambridge, Mass.). Fundamentals of Statistics. Cambridge 38, Mass.: Harvard University Press (38 Quincy St.), 1947. Pp. xvi, 755. Lithotyped. $10.00 (London E.C.4: Oxford University Press [Amen House, Warwick Square], 1948. 55s.) *Reviewed in this issue.*

1639. KEMPTHORNE, O. (Associate Professor of Statistics, Iowa State College, Ames, Iowa). A Simple Approach to Confounding and Fractional Replication in Factorial Experiments. *Biometrika* 34(3–4): 255–72 D '47.*

1640. KENDALL, DAVID G. (University Lecturer in Mathematics, Magdalen College, Oxford University, Oxford, England). On the Generalized "Birth-and-Death" Process. *Ann Math Stat* 19(1) :1–15 Mr '48.*

1641. KENDALL, DAVID G. A Review of Some Recent Work on Discontinuous

Markoff Processes With Applications to Biology, Physics and Actuarial Science. *J Royal Stat Soc* 110(2):130–7 pt 2 '47.*

1642. KENDALL, M. G. (Joint Assistant General Manager and Statistician, Chamber of Shipping of the United Kingdom, Bury Court, St. Mary Axe, London E.C.3, England). **On the Analysis of Oscillatory Time-Series.** *J Royal Stat Soc* 108 (1–2):93–129, discussion 129–41 '45.*

1643. KENDALL, M. G. **The Variance of τ When Both Rankings Contain Ties.** *Biometrika* 34(3–4):297–8 D '47.*

1644. KENNEDY, CLIFFORD W. (Quality Control Engineer, Federal Products Corp., Providence 1, R. I.). **Sensible Standardizations With Statistical Quality Control.** *Prod Eng & Mgmt Mag* 18(6):75–8+ D '46.*

1645. ★KERRICH, J. E. (Senior Lecturer in Mathematics, University of the Witwatersrand, Johannesburg, South Africa). **An Experimental Introduction to the Theory of Probability.** Copenhagen, Denmark: Einar Munksgaard, 1946. Pp. 98.* *To be reviewed.*

1646. KIMBALL, BRADFORD F. (Senior Statistician, State Department of Public Service, 233 Broadway, New York 7, N. Y.). **Some Basic Theorems for Developing Tests of Fit for the Case of the Non-Parametric Probability Distribution Function I.** *Ann Math Stat* 18(4):540–8 D '47.*

1647. KIMBALL, BRADFORD F. **A System of Life Tables for Physical Property Based on the Truncated Normal Distribution.** *Econometrica* 15(4):342–60 O '47.*

1648. KINCAID, W. M. (Instructor in Mathematics, University of Michigan, Ann Arbor, Mich.). **Note on the Error in Interpolation of a Function of Two Independent Variables.** *Ann Math Stat* 19(1):85–8 Mr '48.*

1649. KIRCHEN, CALVIN J. (Statistical Engineer, Technical Department, Remington Arms Co., Bridgeport, Conn.). **Some Useful Statistical Methods.** *Ind Qual Control* 4(2):18–21 S '47.*

1650. KNOWLER, LLOYD A. (Associate Professor of Mathematics and Chairman of the Department, State University of Iowa, Iowa City, Iowa). **Why Statistical Quality Control Gets Results,** pp. 167–70. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1651. KNUDSEN, LILA F. (Statistician, Division of Pharmacology and Food Division, Food and Drug Administration, Federal Security Agency, Washington, D. C.). **The Use of Statistics in Biological Experimentation and Assay.** *J. Assn Offic Agric Chem* 28(4):806–13 N 15 '45.*

1652. KOSSACK, CARL F. (Associate Professor of Mathematics, Purdue University, Lafayette, Ind.). **The Existence of Collectives in Abstract Space.** *Sankhyā* 8(3): 219–34 O '47.*

1653. KRAMER, H. H. (Associate in Agronomy, Agricultural Experiment Station, Purdue University, Lafayette, Ind.), and BURNHAM, C. R. (Professor of Plant Genetics, Agricultural Experiment Station, University of Minnesota, St. Paul, Minn.). **Methods of Combining Linkage Intensity Values from Backcross F_2 and F_3 Genetic Data.** *Genetics* 32(4):379–90 Jl '47.*

1654. KULLBACK, S. (Research Analyst, Army Security Agency, Department of the Army; Lecturer in Statistics, George Washington University; Washington, D.C.). **On the Charlier Type B Series.** *Ann Math Stat* 18(4):574–81 D '47.*

1655. ★LANDAU, HYMAN G. (Mathematician, Ballistic Research Laboratories, Aberdeen Proving Grounds, Md.). **On the Relations Between Certain Criteria for the Estimation of Statistical Parameters.** Doctor's thesis, University of Pittsburgh (Pittsburgh, Pa.), 1946.

1656. LAZARSFELD, PAUL F. (Professor of Sociology, Columbia University, New York, N. Y.). **Factor Analyses of Qualitative Attributes.** Abstract. *Am Psychol* 2(8): 306 Ag '47.*

1657. ★LEHMANN, ERICH L. (Assistant Professor of Mathematics, University of California, Berkeley, Calif.). **Optimum Tests of a Certain Class of Hypotheses Specifying the Value of a Correlation Coefficient.** Doctor's thesis, University of California (Berkeley, Calif.), 1946.

1658. LEHMANN, E. L., and SCHEFFÉ, HENRY (Associate Professor of Engineering, University of California, Los Angeles, Calif.). **On the Problem of Similar Regions.** Abstract. *Ann Math Stat* 18(4):609 D '47.* Also in *B Am Math Soc* 53(11):1129–30 N '47.*

1659. LEHMANN, E. L., and STEIN, C. M. **Most Powerful Tests of Composite Hypotheses: I, Normal Distributions.** Abstract. *Ann Math Stat* 19(1):116–7 Mr '48.*

1660. ★LEMME, MAURICE M. (Assistant Dean of Administration and Associate Professor of Mathematics, Illinois Wesleyan University, Bloomington, Ill.). **On a New Cumulative Frequency Function.** Doctor's thesis, Purdue University (Lafayette, Ind.), 1947.

1661. LEONTIEF, WASSILY (Professor of Economics, Harvard University, Cambridge, Mass.). **Introduction to a The-**

ory of the Internal Structure of Functional Relationships. *Econometrica* 15(4):361–73 O '47.*

1662. LESLIE, R. T. (Section of Mathematical Statistics, Council for Scientific and Industrial Research, South Melbourne S.C.4, Australia). **Over-Estimation of Probable Errors.** Letter. *Nature* 160(4074):751–2 N 29 '47.*

1663. ★LEVENS, A. S. (Associate Professor of Mechanical Engineering, University of California, Berkeley, Calif.). **Nomography.** New York 16: John Wiley & Sons, Inc. (440 Fourth Ave.), 1948. Pp. viii, 176. $3.00.* *To be reviewed.*

1664. LEVINSON, NORMAN (Associate Professor of Mathematics, Massachusetts Institute of Technology, Cambridge, Mass.). **A Heuristic Exposition of Wiener's Mathematical Theory of Prediction and Filtering.** *J Math & Physics* 26(2):110–9 Jl '47.*

1665. LEWIS, WYATT H (Quality Control Engineer, General Electric Co., Ontario, Calif.). **Reestablishing Operator Responsibility for Quality,** pp. 233–49. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1666. LICKLIDER, J. C. R. (Lecturer on Psychology, Harvard University, Cambridge, Mass.), and DZENDOLET, E. (California Institute of Technology, Pasadena, Calif.). **Oscillographic Scatterplots Illustrating Various Degrees of Correlation.** *Sci* 107(2770):121–4 Ja 30 '48.*

1667. LINDBLOM, SVEN G. (Mathematician, Life Insurance Company Folket, Stockholm, Sweden). **On the Connection Between Tests of Significance for Correlation Coefficients and for Differences Between Means.** *Skandinavisk Aktuarietskrift* 29(1–2):12–29 '46.*

1668. LINDLEY, D. V. (National Physical Laboratory, Teddington, Middlesex, England). **Regression Lines and the Linear Functional Relationship.** *Sup J Royal Stat Soc* 9(2):218–44 '47.*

1669. LLOYD, B. H. (Research Engineer, Ontario Research Foundation, 43 Queen's Park, Toronto, Canada). **Statistical Quality Control as Applied in Metal Cutting Plant.** *Mfg & Ind Eng* 25(7):17–8+ Jl '47.*

1670. LOMBARD, HERBERT L. (Director, Division of Cancer and Other Chronic Diseases, Massachusetts Department of Public Health; and Assistant Professor of Hygiene and Public Health, Tufts College Dental School; Boston, Mass.), and DOERING, CARL R. (Assistant Professor of Biostatistics, Harvard School of Public Health, Cambridge, Mass.). **Treatment of the Fourfold Table by Partial Association and Partial Correlation as It Relates to Public Health Problems.** *Biometrics* 3(3):123–8 S '47.*

1671. LUCIA, ESCHSCHOLTZIA L. (Consultant in Statistical Research, California State Department of Public Health, San Francisco 2, Calif.). **Teaching of Public Health Statistics.** *Am J Pub Health* 36(7): 734–8 Jl '46.*

1672. LYMAN, W. J. (Engineering Assistant to Manager, Planning and Development Department, Duquesne Light Co., Duquesne, Iowa). **Service Continuity Standards Aid Economical Distribution Planning.** *Edison Elec Inst B* 15(9):325–9+ S '47.*

1673. ★McCARTHY, PHILIP J. (Assistant Professor of Sociology, Cornell University, Ithaca, N. Y.). **Approximate Solutions for Means and Variances in a Certain Class of Box Problems.** Doctor's thesis, Princeton University (Princeton, N. J.), 1947.

1674. MACCOBY, ELEANOR E. (Graduate Student), and GOODMAN, ROE (Head, Sampling Section, Survey Research Center). (University of Michigan, Ann Arbor, Mich.) **Explanation of the Sampling Method [Used in Survey of Consumer Finances].** *Fed Res B* 33(6):658–63 Je '47.*

1675. McEWEN, G. F. (Professor and Curator of Physical Oceanography, Scripps Institution of Oceanography, University of California, La Jolla, Calif.). **Evaluation of an Integral Occurring in a Fluctuation Problem of Cosmic-Ray Showers.** Abstract. *B Am Math Soc* 54(1):84–5 Ja '48.*

1676. MACEY, H. H. (British Refractories Research Association, Mellor Laboratories, Shelton, Stoke-on-Trent, England). **Some Uses of Statistical Methods in the Heavy Clay Industry.** *Brit Ceramic Soc Trans* 46(6):220–46 Ag '47.* Reprinted from *Brit Refractories Res Assn B* Je '45.

1677. McVAY, FRANCIS E. (Assistant Professor of Economics, North Carolina State College, Raleigh, N. C.). **Sampling Methods Applied to Estimating Numbers of Commercial Orchards in a Commercial Peach Area.** *J Am Stat Assn* 42(240):533–40 D '47.*

1678. MAIZELS, JOAN (Industrial Relations Adviser, Industrial Welfare Society, Inc., 14 Hobart Place, London S.W.1, England), and MAIZELS, ALFRED (Principal, Statistics Division, Board of Trade, Romney House, Tufton St., London S.W.1, England). **The Application of Some Statistical Techniques to Personnel Management.** *Labour Mgmt* 27(281, 283):82–5, 131–3

Ag–S '45, D–Ja '45–46; 28(284, 285):11–15, 32–8 F–Mr, Ap–My '46.*

1679. MALLIK, A. K.; SATAKOPAN, V.; and RAO, S. GOPAL. (Agricultural Meteorology Section, Meteorological Office, Poona, India.) A Study of the Estimation of the Yield of Wheat by Sampling. *Indian J Agric Sci* 15(4):219–25 Ag '45.*

1680. MANUELE, JOSEPH (Director of Quality Control, Westinghouse Electric and Manufacturing Co., East Pittsburgh, Pa.). The Use of Instruments for Controlling Quality. *Machinery* (*New York*) 53 (6):156–9 F '47.*

1681. MANUELE, JOSEPH, and GOFFMAN, CASPER. Statistical Tools for Controlling Quality. *Trans Am Inst Elec Eng* 64: 524–8, discussion 949–51 '46.*

1682. MARCUSE, SOPHIE (Statistician, Bureau of Nutrition and Home Economics, Agricultural Research Administration, U. S. Department of Agriculture, Washington, D. C.). Applying Control Chart Methods to Taste Testing. *Food Ind* 19(3):316–8 Mr '47.*

1683. MARSCHAK, JACOB (Professor of Economics, and Research Director of the Cowles Commission for Research in Economics, University of Chicago, Chicago, Ill.). Statistical Inference From Nonexperimental Observations. Abstract. *Econometrica* 16(1):53–5 Ja '48.*

1684. MATÉRN, BERTIL (State Institute of Forestry Research, Stockholm, Sweden). Metoder att uppskatta noggrannheten vid linje- och provytetaxering. [Methods of Estimating the Accuracy of Line and Sample Plot Surveys.] Includes an English summary, pp. 118–36. Stockholm, Sweden: Statens Skogsforskningsinstitut, 1947. Pp. 138. Paper.*

1685. MAYHEW, W. L. (Anglo-Ceylon General Trusts Co., Ltd., Port Louis, Mauritius). Actuarial Examinations. *J Inst Actuaries Students' Soc* 7(1):54–5 Jl '47.* Letter in reply to 1629.

1686. MAYHEW, W. L., and VAJDA, S. An Application of the Theory of Probability to the Examinations of the Institute of Actuaries. *J Inst Actuaries Students' Soc* 6(2):67–75 O '46.*

1687. MERRELL, MARGARET (Associate Professor of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.). Time-Specific Life Tables Contrasted With Observed Survivorship. *Biometrics* 3(3):129–36 S '47.*

1688. MICHEL, J. G. L. (Physics Department, University of Manchester, Manchester 13, England). Central Difference Formulae Obtained by Means of Operator Expansions. *J Inst Actuaries* 72(335):470–80 Je '46.*

1689. ★MISES, RICHARD VON (Gordon McKay Professor of Aerodynamics and Applied Mathematics, Harvard University, Cambridge, Mass.). Lecture Notes on Mathematical Theory of Probability and Statistics. Harvard University, Graduate School of Engineering Special Publication No. 1. Cambridge 38, Mass.: Harvard Graduate School of Engineering (c/o Miss Natalie Nicholson, Librarian, Pierce Hall), 1947. Pp. 300. Paper, mimeographed. Out of print. Request for review copy refused.

1690. MITCHELL, JAMES A. (Supervisor of Quality Control and Development, Cellulose Esters Division, Tennessee Eastman Corp., Kingsport, Tenn.). Control of the Accuracy and Precision of Industrial Tests and Analyses. *Ind & Eng Chem, Anal Ed* 19(12):961–7 D '47.*

1691. MITCHELL, JAMES A. Quality Control and Chemical Specification, pp. 255–8. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1692. MITCHELL, JAMES A. Quality Control in the Chemical Industry: VI, Quality Control in Raw Materials Acceptance and Chemical Specifications. *Ind Qual Control* 4(3):16–20 N '47.*

1693. MOORE, THOMAS V. (Professor of Psychology); STAFFORD, JOHN W. (Assistant Professor of Psychology); and HSÜ, EN HSI. (Catholic University of America, Washington, D. C.). Obverse Analysis of Personality. *J Personality* 16(1): 11–48 S '47.*

1694. MORAN, P. A. P. (Institute of Statistics, Oxford University, Oxford, England). On the Method of Paired Comparisons. *Biometrika* 34(3–4):363–5 D '47.*

1695. MORAN, P. A. P. The Random Division of an Interval. *Sup J Royal Stat Soc* 9(1):92–8 '47.*

1696. MORAN, P. A. P. Rank Correlation and Permutation Distributions. *Proc Cambridge Philos Soc* 44(1):142–4 Ja '48.*

1697. MORAN, P. A. P. Some Theorems on Time Series, I. *Biometrika* 34(3–4): 281–91 D '47.*

1698. MORROW, DOROTHY J. (Assistant Professor of Statistics, George Washington University, Washington, D. C.). On the Distribution of the Sums of the Characteristic Roots of a Determinantal Equation. Abstract. *B Am Math Soc* 54(1):75–6 Ja '48.*

1699. MOSTELLER, FREDERICK (Associate Professor of Mathematical Statistics, Department of Social Relations, Harvard University, Cambridge, Mass.).

A *k*-Sample Slippage Test for an Extreme Population. *Ann Math Stat* 19(1):58–65 Mr '48.*

1700. ★MOSTELLER, FREDERICK C. On Some Useful "Inefficient" Statistics. Doctor's thesis, Princeton University (Princeton, N. J.), 1947. *See* 884.

1701. MOVSHIN, JOSEPH (Wagner Electric Corp., 6400 Plymouth Ave., St. Louis 14, Mo.). An Introduction to Sampling. *Ind Qual Control* 4(3):21–4 N '47.*

1702. MUENCH, HUGO (Assistant Dean and Professor of Biostatistics, School of Public Health, Harvard University, Cambridge, Mass.). Statistics in the Planning and Evaluation of Health Practices. *Am J Pub Health* 37(10):1273–6 O '47.*

1703. MYERS, R. J. (Actuarial Consultant, Social Security Administration, Washington 25, D. C.). Laymen Don't Have Logarithmic Eyes: The Case Against the Casual Use of Semi Log Charts. *Am Statistician* 1(2):17 O '47.*

1704. NAIR, K. R. (Statistics Department, University College, University of London, London W.C.1, England). A Note on the Mean Deviation From the Median. *Biometrika* 34(3–4):360–2 D '47.*

1705. NANDA, D. N. (Graduate Student, University of North Carolina, Chapel Hill, N. C.). Certain Results on the Roots of a Determinantal Equation. Abstract. *B Am Math Soc* 54(1):76 Ja '48.*

1706. NANDA, D. N. Distribution of a Root of Determinantal Equation. *Ann Math Stat* 19(1):47–57 Mr '48.* Abstract 18(4):607 D '47.*

1707. NANDA, D. N. Distribution of the Sum of Roots of a Determinantal Equation Under a Certain Condition. Abstract. *Ann Math Stat* 19(1):120 Mr '48.*

1708. NANDI, H. K. (Calcutta University, Calcutta, India). Misuse of Statistics. *Calcutta Stat Assn B* (1):19–21 Ag '47.*

1709. NANDI, H. K. Statistical Organizations in India. *Calcutta Stat Assn B* (2):57–64 N '47.*

1710. NANDI, H. K. Tests of Significance on Multivariate Samples. *Calcutta Stat Assn B* (1):42–3 Ag '47.*

1711. NEYMAN, J. (Professor of Mathematics and Director of the Statistical Laboratory, Univerity of California, Berkeley, Calif.). Outline of Statistical Treatment of the Problem of Diagnosis. *Pub Health Rep* 62(40):1449–56 O 3 '47.* Discussion of 1873.

1712. NISSEN, B. H. (Director, Brewery Division Laboratories, Anheuser-Busch, Inc., St. Louis, Mo.). Sampling Methods, pp. 201–11. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1713. O'CONNOR, H. J. (Technical Assistant, Development Division, Union Carbide & Carbon Research Labs., Inc., Niagara Falls, N. Y.). On Simple Correlation. *Ind Qual Control* 4(2):9–13 S '47.*

1714. OLDS, EDWIN G. (Associate Professor of Mathematics, Carnegie Institute of Technology, Pittsburgh, Pa.). Acceptance Sampling by Variables, pp. 131–44. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1715. OLIFIERS, E. (caixa Postal 8, Petropolis (R-J), Brazil). The Evolution of the Exposure Formulae. Discussion by BERT A. WINTER (Assistant Actuary, Prudential Insurance Co., Newark 1, N. J.), pp. 289–92. *Trans Actuarial Soc Am* 48 (117, 118):76–94, 289–95 My, O '47.*

1716. ★ORCUTT, GUY HENDERSON (Senior Research Worker, Department of Applied Economics, Cambridge University, Cambridge, England). Statistical Methods and Tools for Finding Natural Laws in the Field of Economics. Doctor's thesis, University of Michigan (Ann Arbor, Mich.), 1945.

1717. OWENS, WILLIAM A., JR. (Professor of Psychology, Iowa State College, Ames, Iowa). An Empirical Study of the Relationship Between Item Validity and Internal Consistency. *Ed & Psychol Meas* 7(2):281–8 su '47.*

1718. PARKIN, GUY G. (Director, Industrial and Statistical Control, Minnesota Mining and Manufacturing Co., 900 Farquier Ave., St Paul 6, Minn.). The application of Statistical Quality Control to the Control of Waste and the Measure of Personnel Efficiency, pp 193–9. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1719. PAULSON, EDWARD (Instructor in Mathematical Statistics, The University of North Carolina, Chapel Hill, N. C.), and WALLIS, W. ALLEN. Chap. 7, Planning and Analyzing Experiments for Comparing Two Percentages, pp. 247–65. In *Selected Techniques of Statistical Analysis* (*see* 1512).

1720. PEACH, PAUL (Associate Professor, Institute of Statistics, North Carolina State College, Raleigh, N. C.). Teaching Statistics to Engineering Students, pp. 251–3. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1721. PERCIVAL, D. A. (Superintendent of Census, Nicosia, Cyprus). An Experiment in Census Tabulation: Census of Cyprus, 1946. *J Royal Stat Soc* 110(2): 138–43 pt 2 '47.*

1722. PERKS, WILFRED (Assistant Actuary, Pearl Assurance Co., Ltd., Holborn, London W.C.1, England). A Simple Proof of Gauss's Inequality. *J Inst Actuaries Students' Soc* 7(1):38–41 Jl '47.*

1723. PERKS, WILFRED. Two-Variable Developments of the *n*-Ages Method. *J Inst Actuaries* 72(335):377–97, discussion 397–414 Je '46.*

1724. PLACKETT, R. L. (Lecturer in Statistics, University of Liverpool, Liverpool, England). Cyclic Intrablock Subgroups and Allied Designs. *Sankhyā* 8(3): 275–6 O '47.*

1725. PLACKETT, R. L. An Exact Test for the Equality of Variances. *Biometrika* 34(3–4):311–9 D '47.*

1726. POLITZ, ALFRED (President), and BRUMBACH, RICHARD (Field Supervisor). (Alfred Politz Research, Inc., 400 Madison Ave., New York 17, N. Y.) Can an Advertiser Believe What Mail Surveys Tell Him? *Ptr Ink* 219(12):48+ Je 20 '47.*

1727. PRUTHI, HEM SINGH, and PRADHAN, S. (Imperial Agricultural Research Institute, New Delhi, India.) Methods of Computing Pest Incidence. *Indian J Agric Sci* 15(5):265–9 O '45.*

1728. QUENOUILLE, M. H. (Lecturer in Statistics and Head of the Department, Marischal College, Aberdeen University, Aberdeen, Scotland). A Large-Sample Test for the Goodness of Fit of Autoregressive Schemes. *J Royal Stat Soc* 110(2):123–9 pt 2 '47.*

1729. QUENOUILLE, M. H. Notes on the Calculation of Autocorrelations of Linear Autoregressive Schemes. *Biometrika* 34(3–4):365–7 D '47.*

1730. QUENOUILLE, M. H. On the Problem of Random Flights. *Proc Cambridge Philos Soc* 43(4):581–2 O '47.*

1731. QUENSEL, CARL-ERIK (Director of the Institute of Statistics and Professor of Statistics, University of Lund, Lund, Sweden). Studies of the Logarithmic Normal Curve. *Skandinavisk Aktuarietskrift* 28(3–4):141–53 '45.*

1732. QUENSEL, CARL-ERIK. The Validity of the z-Criterion When the Variates Are Taken From Different Normal Populations. *Skandinavisk Aktuarietskrift* 30(1–2):44–55 '46.*

1733. RAO, C. RADHAKRISHNA. (King's College, University of Cambridge, Cambridge, England). Factorial Experiments from Combinatorial Arrangements of Arrays. *Sup J Royal Stat Soc* 9(1):128–39 '47.*

1734. RAO, C. RADHAKRISHNA. General Methods of Analysis for Incomplete Block Designs. *J Am Stat Assn* 42(240):541–61 D '47.*

1735. RAO, C. RADHAKRISHNA. Large Sample Tests of Statistical Hypotheses Concerning Several Parameters With Applications to Problems of Estimation. *Proc Cambridge Philos Soc* 44(1):50–7 Ja '48.*

1736. RAO, C. RADHAKRISHNA. A Statistical Criterion to Determine the Group to Which an Individual Belongs. Letter. *Nature* 160(4076):835–6 D 13 '47.*

1737. RAO, W. V. B. SUNDARO, and REDDY, M. K. (Imperial Agricultural Research Institute, New Delhi, India.) Studies on the Methods of Estimating Total Bacterial Counts in the Soil. *Indian J Agric Sci* 15(3):111–5 Je '45.*

1738. REIERSÖL, OLAV (Teacher of Mathematical Statistics, University of Oslo, Oslo, Norway). Residual Variables in Regression and Confluence Analysis. *Skandinavisk Aktuarietskrift* 28(3–4):201–17 '45.*

1739. RICE, S. O. (Member of Technical Staff, Bell Telephone Laboratories, Inc., 463 West St., New York 14, N. Y.). Statistical Properties of a Sine Wave Plus Random Noise. *Bell System Tech J* 27: 109–57 Ja '48.*

1740. RICH, WILLIS H. (Professor of Biology, Stanford University, Stanford, Calif.), and TERRY, M. C. (Physician, U. S. Veterans Administration, retired, Palo Alto, Calif.). The Industrial "Control Chart" Applied to the Study of Epidemics. *Pub Health Rep* 61(42):1501–11 O 18 '46.*

1741. RICKER, WILLIAM E. (Associate Professor of Zoology, Indiana University, Bloomington, Ind.). A Method of Estimating Minimum Size Limits for Obtaining Maximum Yield. *Copeia* (2):84–94 Je 30 '45.*

1742. ROBBINS, HERBERT (Associate Professor of Mathematical Statistics, University of North Carolina, Chapel Hill, N. C.). Convergence of Distributions. *Ann Math Stat* 19(1):72–6 Mr '48.*

1743. ROBINSON, DANIEL SOMMER (Professor of Philosophy and Director of the School of Philosophy, University of Southern California, Los Angeles, Calif.). Sect. 6, Probability and the Statistical Methods, pp. 227–50. In his *The Principles of Reasoning:* An Introduction to Logic and Scientific Method, Third Edition. New York 1: D. Appleton-Century Co. (35 West 32nd St.), Inc., 1947. Pp. xx, 410. $3.00.*

1744. ★ROGERS, W. T. (Metallurgical

Statistician and Supervisor of the Statistical Division, Metallurgical Department, National Tube Co., Lorain, Ohio). Multiple Correlation Applied to Steel Plant Problems. Reprint No. 27 from *Trans Am Soc Metals*, Vol. 40, 1948. Chicago, Ill.: American Society for Metals, 1948. Pp. 16. Paper. $0.25.*

1745. ROGERS, W. T. Practical Application of Statistical Methods in a Quality Control Program. *Trans Am Soc Metals* 36:361–82, discussion 382–8 '46.*

1746. ROGERS, W. T. Statistical Methods in Quality Control. Abstract. *Chem & Met Eng* 53(4):274+ Ap '46.*

1747. ROOS, CHARLES F. (President, Econometric Institute, Inc., 500 Fifth Ave., New York 18, N. Y.). Business Planning and Statistical Analysis. Abstract. *Econometrica* 16(1):49–50 Ja '48.* Also in *Comm & Fin Chr* 166(4630):1110+ S 18 '47.*

1748. ROSENTHAL, IRVING (Associate Actuary, Guardian Life Insurance Co., New York 3, N. Y.). Limits of Retention for Ordinary Life Insurance. *Rec Am Inst Actuaries* 36(73):6–22 My '47.*

1749. ROSS, C. C. (Late Professor of Educational Psychology, University of Kentucky, Lexington, Ky.). Chap. 8, Statistical Analysis of Test Results, pp. 216–51; Chap. 9, The Graphical Representation of Educational Data, pp. 252–81. In his *Measurement in Today's Schools*, Second Edition. New York 11: Prentice-Hall, Inc. (70 Fifth Ave.), 1947. Pp. xvi, 551. $4.50.*

1750. ROY, S. N. (Calcutta University, Calcutta, India). Notes on Testing of Composite Hypotheses. *Sankhya* 8(3):257–70 O '47.*

1751. ROY, S. N. A Note on the Relation Between Testing of Hypothesis and Estimation by Confidence Interval. *Calcutta Stat Assn B* (2):89–91 N '47.*

1752. ROY, S. N. Some Aspects of the Organization of Statistical Work in India. *Calcutta Stat Assn B* (2):49–57 N '47.*

1753. Royal Statistical Society, Council. Report on the Teaching of Statistics in Universities and University Colleges. *J Royal Stat Soc* 110(1):51–7 pt 1 '47.*

1754. Royal Statistical Society, Examinations Committee, MAURICE KENDALL, Chairman. First Report of the Examinations Committee, as Amended and Adopted by the Council. *J Royal Stat Soc* 110(1):58–71 pt 1 '47.*

1755. RUBIN, HERMAN (School of Mathematics, Institute for Advanced Study, Princeton, N. J.). The Asymptotic Analogue of the Theorem of Cramér and Rao. Abstract. *Ann Math Stat* 19(1):121 Mr '48.*

1756. RUBIN, HERMAN. Some Results on the Distribution of Quadratic Forms From Gaussian Stochastic Processes: Preliminary Report. Abstract. *Ann Math Stat* 18(4):610 D '47;* *B Am Math Soc* 53(11):1130 N '47.*

1757. RUTHERFORD, C. D. (Associate Actuary, Sun Life Assurance Company of Canada, Montreal, Canada). Actuarial Mathematics and Statistics: Symposium on Statistics, II, pp. 25–30. In *Proceedings of the First Canadian Mathematical Congress*, Montreal, 1945. Toronto, Canada: University of Toronto Press, 1946. Pp. xliv, 367. $3.25.*

1758. ★SAMUELSON, PAUL ANTHONY (Professor of Economics, Massachusetts Institute of Technology, Cambridge, Mass.). Foundation of Economic Analysis. Harvard Economic Studies, Vol. 80. Cambridge 38, Mass.: Harvard University Press (38 Quincy St.), 1947. Pp. xii, 447. $7.50.* *To be reviewed.*

1759. SAPPENFIELD, BERT R. (Assistant Professor of Psychology, Montana State University, Missoula, Mont.). A Rapid Method of Computing Standard Scores. *J Appl Psychol* 31(6):638–9 D '47.*

1760. ★SCARBOROUGH, J. B. (Professor of Mathematics, United States Naval Academy, Annapolis, Md.), and WAGNER, R. W. (Assistant Professor of Mathematics, Oberlin College, Oberlin, Ohio). Fundamentals of Statistics. Boston 17, Mass.: Ginn and Co. (Statler Building), 1948. Pp. vii, 145. $2.40.* *To be reviewed.*

1761. SCHROCK, EDWARD M. (Quality Control Engineer, Refrigerator Quality Control Division, General Electric Co., Erie, Pa.). Using the Quality Control Chart to Identify Assignable Causes, pp. 97–101. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1762. SEAL, H. L. (Head of Statistical Branch, Admiralty, Whitehall, London S.W.1, England). Gambler's Ruin. *J Inst Actuaries Students' Soc* 6(1):35–7 Jl '46.*

1763. SEAL, H. L. A Historical Note on the Use of χ^2 to Test the Applicability of a Mortality Table Graduation. *J Inst Actuaries Students' Soc* 6(4):185–7 Ap '47.*

1764. SEAL, H. L. Multiple Decrement Tables and the Force of Mortality—A Historical Note. *J Inst Actuaries Students' Soc* 6(4):197–9 Ap '47.*

1765. SEAL, H. L. A Probability Distribution of Deaths at Age x When Policies Are Counted Instead of Lives. *Skandinavisk Aktuarietidskrift* 30(1–2):18–43 '47.*

1766. SHEA, H. GREGORY (Associate Editor, Electronic Industries, 480 Lexing-

ton Ave., New York 17, N. Y.). Meeting Specs in UHF. *Electronic Ind* 4(5):84–5+ My '45.*

1767. SHEAFFER, C. R. (President, W. A. Sheaffer Pen Co., Fort Madison, Iowa). **What Top Management Expects From Quality Control,** pp. 1–6. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1768. SILBERSTEIN, LUDWIK (*Lately of* Kodak Research Laboratories, Eastman Kodak Co., Rochester, N. Y.). **On Two Accessories of Three-Dimensional Colorimetry: I, The Probable Error of Colorimetric Tensor Components as Derived From a Number of Color Matchings; II, The Determination of the Principal Colorimetric Axes at Any Point of the Color Threefold.** *J Opt Soc Am* 36(8):464–8 Ag '46.*

1769. SIMON, LESLIE E. (Colonel, Ordnance Department, and Director, Ballistic Research Laboratories, Aberdeen Proving Ground, Md.). **Dollars for Your Thoughts: How the Ordnance Department Used Statistical Methods of Quality Control to Prevent Destruction of Tank Armor During Tests and Found Huge Savings in Materials and Costs.** *Ind Stand* 17(8): 203–6 Ag '46.*

1770. SIMON, LESLIE E. **Sampling and Sorting Ammunition for the Attack on Normandy,** pp. 275–81. In *Conference Papers: First Annual Convention of America Society for Quality Control . . .* (*see* 1380).

1771. SIMONSEN, W. (Professor of Insurance Mathematics, University of Copenhagen, Copenhagen, Denmark). **On Distributions of Functions of Samples From a Normally Distributed Infinite Population.** *Skandinavisk Aktuarietskrift* 28(1–2):20–43 '45.*

1772. SLATER, PATRICK (Principal Research Officer at The Social Survey, 39–41 Nottingham Place, London N.W.1, England). **The Factor Analysis of a Matrix of 2×2 Tables.** *Sup J Royal Stat Soc* 9(1): 114–27 '47.*

1773. SMALLWOOD, HUGH M. (Research and Technical Development Department, U. S. Rubber Co., Passaic, N. J.). **Design of Experiments in Industrial Research.** *Ind & Eng Chem, Anal Ed* 19(12): 950–2 D '47.*

1774. ★SMART, L. EDWIN (Professor of Economics), and ARNOLD, SAM (Assistant Professor in Business Research). (Ohio State University, Columbus, Ohio.) **Practical Rules for Graphic Presentation of Business Statistics.** Columbus, Ohio: Bureau of Business Research, College of Commerce and Administration, Ohio State University, 1947. Pp. xii, 86, $2.00.* *To be reviewed.*

1775. SMITH, GEORGE C., JR. (Regional Economist, U. S. Department of Commerce, Washington, D. C.). **Lorenz Curve Analysis of Industrial Decentralization.** *J Am Stat Assn* 42(240):591–6 D '47.*

1776. SMITH, GEORGE F. (President, Johnson & Johnson, New Brunswick, N. J.). **A Management Viewpoint of Quality Control,** pp. 7–10. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1776a. SMITH, GEORGE F. **The Management Viewpoint: Symposium on Statistical Methods in Experimental and Industrial Chemistry.** *Ind & Eng Chem, Anal Ed* 19(12):943–6 D '47.*

1777. SNYDER, LAURENCE H. (Dean, Graduate College, Professor of Zoology; and Professor of Medical Genetics, College of Medicine; University of Oklahoma, Norman, Okla.). **The Principles of Gene Distribution in Human Populations.** *Yale J Biol & Med* 19(5):817–33 My '47.*

1778. SPEAKMAN, A. J. **Statistical Methods of Maintenance Control.** *Mech World* 122(3175):521–3 N 21 '47.*

1779. SPENCER-SMITH, J. L. (Senior Assistant in Spinning Department, Linen Industry Research Association, Lambeg, Antrim, Northern Ireland). **The Estimation of Fibre Quality.** *J Textile Inst* 38(8): P257–71, discussion P271–2 Ag '47.*

1780. SPENCER-SMITH, J. L. **The Oscillatory Properties of the Moving Average.** *Sup J Royal Stat Soc* 9(1):104–13 '47.*

1781. STEEN, J. R. (Director of Quality Control, Sylvania Electric Products, Inc., Flushing, N. Y.). **The Use of Modified Limits in Process Quality Control,** pp. 103–10. In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1782. STEIN, CHARLES (Lecturer in Mathematics, University of California, Berkeley, Calif.). **On Sequences of Experiments.** Abstract. *Ann Math Stat* 19(1): 117–8 Mr '48.*

1783. STEINER, E. H. (The British Food Manufacturing Industries Research Association, 2–4 Dalmeny Ave., London N.7, England). **Application of Statistical Methods in Calculating Proportions of Ingredients in Certain Food Products: The Application of Statistical Methods to Food Problems.** *Analyst* 73(862):15–29 Ja '48.* Abstract *Chem & Ind* (3):43–4 Ja 18 '47.*

1784. STEINHAUS, HUGO (Professor, University of Warsaw, Warsaw, Poland). **The Problem of Fair Division.** Abstract. *Econometrica* 16(1):101–4 Ja '48.*

1785. STEPHAN, FREDERICK F. (Director of Research Program in Education and Professor of Sociology, Princeton University, Princeton, N. J.). History of the Uses of Modern Sampling Procedures. Abstract. *Econometrica* 16(1):72–4 Ja '48.*

1786. STEPHAN, FREDERICK F., and McCARTHY, PHILIP J. Sampling Opinion, Attitudes, and Consumer Wants. Reprinted from *Items*, June 1947. *Am Statistician* 1(3):6–7 D '47.*

1787. STONE, J. R. N. (Director, Department of Applied Economics, University of Cambridge, Cambridge, England). Prediction From Autoregression Schemes and Linear Stochastic Difference Systems. Abstract. *Econometrica* 16(1):38–9 Ja '48.*

1788. SUKHATME, P. V. (Statistical Adviser, Statistical Branch, Indian Council of Agricultural Research, New Delhi, India). The Indian Society of Agricultural Statistics. *Cur Sci* 17(1):11 Ja '48.*

1789. SUKHATME, P. V. Random Sampling for Estimating Rice Yield in Madras Province. *Indian J Agric Sci* 15(6): 308–18 D '45.*

1790. SUKHATME, P. V. Use of Small-Size Plots in Yield Surveys. *Nature* 160(4068):542 O 18 '47.*

1791. SULLIVAN, J. J. (Quality Control Director, Reliance Manufacturing Co., 212 W. Monroe St., Chicago 6, Ill.). Practical Application of Quality Control to the Problems of Garment Manufacturing, pp. 271–4. In *Conference Papers: First Annual Convention of American Society for Quality Control* . . . (*see* 1380).

1792. ★SUR, A. K. (Editor, Calcutta Stock Exchange Official Year Book). Statistics: How to Handle Them: A Practical Handbook for Students and Businessmen With a Special Chapter on Business Forecasting. Bombay, India: D. B. Taraporevala Sons & Co. (210 Hornby Road), [1947?]. Pp. ix, 121. Rs. 3–15.*

1793. SWANSON, C. O. (Professor of Petrology, University of British Columbia, Vancouver, B. C., Canada). Probabilities in Estimating the Grade of Gold Deposits. *Trans Can Inst Min & Metall* 48:323–50 '45.*

1794. SWINEFORD, FRANCES (Research Associate, Department of Education), and HOLZINGER, KARL J. (Professor of Education). (University of Chicago. Chicago, Ill.) Selected References on Statistics, the Theory of Test Construction, and Factor Analysis. *Sch R* 55(6):363–8 Je '47.*

1795. TAYLOR, CHARLES R. (Supervising Metallurgist, Research Laboratories, American Rolling Mill Co., Middletown, Ohio). Graphical Correlation: Its Application to Steel Mill Problems. *Iron Age* 160(19):78–84 N 6 '47.*

1796. TAYLOR, J. J. (Chief Engineer, Insulator Division, Ohio Brass Co., Columbus, Ohio). Statistical Methods Applied to Insulator Development and Manufacture. *Trans Am Inst Elec Eng* 64:495–9, discussion 952 '45.*

1797. TAYLOR, W. S. (Supervisor of the Resin Department, Hercules Powder Co., Mansfield, Mass.). Use of Statistical Methods for Time Study of Batch Processes. *Ind Qual Control* 4(4):18–22 Ja '48.*

1798. THOMSON, D. HALTON (Engineer, The Portsmouth Water Co., 26–28 Commercial Road, Portsmouth, England). Approximate Formulae for the Percentage of the Incomplete Beta Function and of the χ^2 Distribution. *Biometrika* 34(3–4):368–72 D '47.*

1799. THOMSON, GODFREY (Professor of Education, University of Edinburgh, Edinburgh 8, Scotland). The Maximum Correlation of Two Weighted Batteries (Hotelling's "Most Predictable Criterion"). *Brit J Psychol, Stat Sect* 1(1): 27–34 O '47.*

1800. THORNDIKE, ROBERT L. (Associate Professor of Education, Teachers College, Columbia University, New York, N. Y.). Logical Dilemmas in the Estimation of Reliability, pp. 21–30, discussion 31–40. In *National Projects in Educational Measurement:* A Report of the 1946 Invitational Conference on Testing Problems. Edited by K. W. Vaughn. American Council on Education Studies, Vol. 11; Series 1, Reports of Committees and Conferences, No. 28. Washington 6, D. C.: American Council on Education (744 Jackson Place), August 1947. Pp. vii, 80. Paper. $1.00.*

1801. ★THORNDIKE, ROBERT L. Research Problems and Techniques. Army Air Forces, Aviation Psychology Program Research Reports, No. 3. Washington, D. C.: Government Printing Office, 1947. Pp. viii, 163. $1.00.*

1802. THOULESS, ROBERT H. (Reader in Educational Psychology, Cambridge University, Cambridge, England). A Method of Correcting for Dice Bias in Evaluating PK Test Data. Letter. Reprinted from *J Soc Psychical Res*. *J Parapsychol* 11(3):231–5 S '47.*

1803. THURSTONE, L. L. (Charles F. Grey Distinguished Service Professor of Psychology, University of Chicago, Chicago 37, Ill.). Factorial Analysis of Body Measurements. *Am J Physical Anthropol* 5(1):15–28 Mr '47.*

1804. TIFFIN, JOSEPH (Professor of

Psychology, Purdue University, Lafayette, Ind.). **Appendix A, Elementary Statistical Procedures, pp. 485–520.** In his *Industrial Psychology*, Second Edition. New York 11: Prentice-Hall, Inc. (70 Fifth Ave.), 1947. Pp. xxi, 553. $4.00, school edition; $5.35, trade edition.*

1805. TILDESLEY, M. L. **Choice of the Unit of Measurement in Anthropometry.** *Man* 47:72–8 My '47.*

1806. TINTNER, GERHARD (Professor of Mathematics and Economics, and Research Professor of Agricultural Economics, Iowa State College, Ames, Iowa). **Applications of Carnap's Probability Theory to Statistical Inference. Abstract.** *Ann Math Stat* 19(1):120–1 Mr '48.*

1807. TIPPETT, L. H. C. (Statistician to the British Cotton Industry Research Association, Shirley Institute, Didsbury, Manchester, England). **Quality Control in the Textile Industry.** *Textile Inst J* 38(12): P594–8, discussion P598–600 D '47.*

1808. TIPPETT, L. H. C. **The Study of Industrial Efficiency, With Special Reference to the Cotton Industry.** *J Royal Stat Soc* 110(2):108–16, discussion 116–22 pt 2 '47.*

1809. TOOPS, HERBERT A. (Professor of Psychology, Ohio State University, Columbus, Ohio). **The Use of Addends in Experimental Control, Social Census, and Managerial Research.** *Psychol B* 45(1): 41–74 Ja '48.*

1810. TÖRNQVIST, LEO (Assistant Professor of Statistics, University of Helsinki, Helsinki, Finland). **On the Distribution Function for a Function of n Statistic Variables and the Central Limit Theorem in the Mathematical Theory of Probability.** *Skandinavisk Aktuarietskrift* 29(3–4):206–29 '46.*

1811. TÖRNQVIST, LEO. **On the Economic Theory of Lottery-Gambles.** *Skandinavisk Aktuarietskrift* 28(3–4):228–46 '45.*

1812. TOULOUSE, J. H. (Chief Engineer, Quality and Specifications Department, Owen-Illinois Glass Co., Toledo 1, Ohio). **Statistical Control.** *Glass Ind* 29(2): 79–80+ F '48.*

1813. ★TSAO, FEI (Professor of Educational Psychology, National Central University, Chungking, China). **General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in Subclasses.** Doctor's thesis, University of Minnesota (Minneapolis, Minn.), 1946.

1814. TUCKER, JOHN, JR. (Chief, Concreting Materials Section, National Bureau of Standards, Washington, D. C.). **Effect of Dimensions of Specimens Upon the Precision of Strength Data.** *Am Soc Testing Mat Proc* 45:952–9, discussion 960 '45.*

1815. TUCKER, JOHN, JR. **The Maximum Stresses Present at Failure of Brittle Materials.** *Am Soc Testing Mat Proc* 45: 961–73, discussion 974–5 '45.*

1816. TUKEY, JOHN W. (Assistant Professor of Mathematics, Princeton University, Princeton, N. J.; and Bell Telephone Laboratories, Murray Hill, N. J.). **Approximate Weights.** *Ann Math Stat* 19(1):91–2 Mr '48.*

1817. TUKEY, JOHN W. **Non-Parametric Estimation: II, Statistically Equivalent Blocks and Tolerance Regions—The Continuous Case.** *Ann Math Stat* 18(4): 529–39 D '47.*

1818. TUKEY, JOHN W. **Nonparametric Estimation: III, Statistically Equivalent Blocks and Multivariate Tolerance Regions—The Discontinuous Case.** *Ann Math Stat* 19(1):30–9 Mr '48.*

1819. TYLER, LEONA E. (Associate Professor of Psychology, University of Oregon, Eugene, Oregon). **Chap. 10, Logic and Statistical Concepts, pp 241–55.** In her *The Psychology of Human Differences.* New York 1: D. Appleton-Century Co., Inc. (35 West 32nd St.), 1947. Pp. xiii, 420. $3.75.*

1820. VAJDA, S. (Statistician, Statistical Branch, Admiralty, London S.W.1, England). **Statistical Investigation of Casualties Suffered by Certain Types of Vessels.** *Sup J Royal Stat Soc* 9(2):141–63, discussion 164–75 '47.*

1821. VAJDA, S. **Statistical Methods. A report on a meeting of mathematical statisticians held on September 3–4, 1947 in Gottingen, Germany.** *Nature* 160(4073): 724–5 N 22 '47.*

1822. VAJDA, S. **The Stratified Semi-Stationary Population.** *Biometrika* 34(3–4): 243–54 D '47.*

1823. VAUGHAN, HUBERT (General Secretary and Actuary, Mutual Life and Citizens' Assurance Co., Ltd., Sydney, Australia). **Some Notes on Interpolation.** *J Inst Actuaries* 72(335):482–98 Je '46.*

1824. VILLARS, D. S. (Professor of Chemistry and Head of the Science Department, Jersey City Junior College, Jersey City, N. J.). **A Significance Test and Estimation in the Case of Exponential Regression.** *Ann Math Stat* 18(4):596–600 D '47.*

1825. ★VOTAW, DAVID F., JR. (Instructor in Mathematics, Yale University, New Haven, Conn.). **Testing Compound Symmetry in a Normal Multivariate Distribution.** Doctor's thesis, Princeton University (Princeton, N. J.), 1947. Abstract _Ann Math Stat_ 18(4):608 D '47*; _B Am Math Soc_ 53(11):1130 N '47.*

1826. ★WAGNER, CHARLES C. (Late Professor of Mathematics), JOHNSON, EVAN, JR. (Associate Professor of Mathematics), HAGEN, BEATRICE L. (Associate Professor of Mathematics), (Pennsylvania State College, State College, Pa.); and GRAVES, CLYDE H. (Operations Analyst, Operations Analysis, Air Defense Command, Mitchell Field, N. Y.). Introduction to Statistics, Revised Edition. Ann Arbor, Mich.: Edwards Brothers, Inc., 1946. Pp. iii, 139. Paper, lithotyped. $1.75.*

1827. WALD, ABRAHAM (Professor of Mathematical Statistics and Head of the Department, Columbia University, New York, N. Y.). Asymptotic Properties of the Maximum Likelihood Estimate of an Unknown Parameter of a Discrete Stochastic Process. _Ann Math Stat_ 19(1):40–6 Mr '48.*

1828. WALD, ABRAHAM. An Essentially Complete Class of Admissible Decision Functions. _Ann Math Stat_ 18(4):549–55 D '47.*

1829. WALD, ABRAHAM. Foundations of a General Theory of Sequential Decision Functions. _Econometrica_ 15(4):279–313 O '47.*

1830. WALD, ABRAHAM. A Note on Regression Analysis. _Ann Math Stat_ 18(4):586–9 D '47.*

1831. WALKER, HELEN (Professor of Education, Teachers College, Columbia University, New York, N. Y.). Certain Unsolved Statistical Problems of Importance in Psychological Research. _Harvard Ed R_ 17(4):297–304 fall '47.*

1832. WALLIS, W. ALLEN (Professor of Statistics and Business Economics, The University of Chicago, Chicago, Ill.). Chap. 1, Use of Variables in Acceptance Inspection for Percent Defective, pp. 3–93. In _Selected Techniques of Statistical Analysis_ (_see_ 1512).

1833. WALLIS, W. ALLEN. Chap. 17, Elements of Sequential Analysis, pp. 417–29. In _Selected Techniques of Statistical Analysis_ (_see_ 1512).

1834. WALSH, JOHN E. (Design Specialist, Douglas Aircraft Corp., Santa Monica, Calif.). Loss of Information in t-Tests With Unbalanced Samples: Preliminary Report. Abstract. _Ann Math Stat_ 18(4):611 D '47.*

1835. WALSH, JOHN E. On the Power Efficiency of a t-Test Formed by Pairing Sample Values. _Ann Math Stat_ 18(4):601–4 D '47.*

1836. WALSH, J. E. On the Power Function of the "Best" t-Test Solution of the Behrens-Fisher Problem. Abstract. _Ann Math Stat_ 19(1):117 Mr '48.*

1837. WALSH, JOHN E. On the Use of the Non-Central t-Distribution for Comparing Percentage Points of Normal Populations. _Ann Math Stat_ 19(1):93–4 Mr '48.*

1838. WALSH, JOHN E. Some Significance Tests for the Mean Using the Sample Range and Midrange. Abstract. _Ann Math Stat_ 18(4):607–8 D '47.* Also in _B Am Math Soc_ 53(11):1130–1 N '47.*

1839. WALSH, JOHN E. Some Significance Tests for the Median Which Are Valid Under Very General Conditions: Preliminary Report. Abstract. _Ann Math Stat_ 18(4):610–11 D '47.*

1840. WEAVER, CHALMERS L. (Assistant Actuary, New England Mutual Life Insurance Co., 501 Boylston St., Boston, Mass.). Actuarial Note: Allowance for Rounding Errors in the Summation Check. _Trans Actuarial Soc Am_ 48(118):267–9 O '47.*

1841. WEAVER, C. L. A Simple Analytic Proof of a General χ^2 Theorem. _Am Math Mo_ 54(9):529–33 N '47.*

1842. WECK, FRANK A. (Metropolitan Life Insurance Co., New York 10, N. Y.). The Mortality Rate and Its Derivation From Actual Experience. _Rec Am Inst Actuaries_ 36(73):23–54 My '47.*

1843. WEEKS, H. T. (Cadbury Bros., Ltd., Bournville, Birmingham 30, England). Statistics and the Statistician in Industry. _J Royal Stat Soc_ 110(2):95–8, discussion 98–107 pt 2 '47.*

1844. ★WELD, WALTER E. How to Chart: Facts From Figures With Graphs. Preface by Allan C. Haskell. Norwood, Mass.: Codex Book Co., Inc., 1947. Pp. xiv, 218. $3.00.* _To be reviewed._

1845. WENTWORTH, CHESTER K. (Geologist, Honolulu Board of Water Supply, Honolulu, Hawaii). Cycles in Rainfall and Validity in Prediction of Rainfall in Hawaii. _Pacific Sci_ 1(4):215–20 O '47.*

1846. WERNIMONT, G. (Analytic and Research Chemist, Eastman Kodak Co., Rochester 4, N. Y.). Statistical Control in Chemical Production. Abstract. _Chem Eng_ 55(1):272+ Ja '48.*

1847. WESCOTT, MASON E. (Assistant Professor of Mathematics, Northwestern University, Evanston, Ill.). Attribute Charts in Quality Control, pp. 49–70. In _Conference Papers: First Annual Convention of American Society for Quality Control_ . . . (_see_ 1380).

1848. WESCOTT, MASON E. Measurement Charts in Quality Control, pp. 29–48. In _Conference Papers: First Annual Con-_

vention of American Society for Quality Control . . . (*see* 1380).

1849. **WESTMAN, A. E. R.** (Director, Department of Chemistry, Ontario Research Foundation, 43 Queens Park, Toronto 5, Canada). **Introduction to Statistics.** *ASTM B* (148):61–3 O '47.*

1850. **WESTMAN, A. E. R., and FREEMAN, R. W. S.** (Stamp & Stencil Co., Ltd., 8 George St., Hamilton, Canada). **Statistical Control of the Manufacture of Manganese Steel Tank Shoes.** *Can Metals & Met Ind* 8(6):38–43 Je '45.*

1851. **WHERRY, ROBERT J.** (Professor of Psychology, University of North Carolina, Chapel Hill, N. C.). **Multiple Bi-Serial and Multiple Point Bi-Serial Correlation.** *Psychometrika* 12(3):189–95 S '47.*

1852. **WHITFIELD, J. W.** (Applied Psychology Research Unit, Medical Research Council, Psychological Laboratory, University of Cambridge, Cambridge, England). **Rank Correlation Between Two Variables, One of Which is Ranked, the Other Dichotomous.** *Biometrika* 34(3–4):292–6 D '47.*

1853. **WICKLIFFE, JOHN K.** (Statistical Department, Swift & Co., Chicago, Ill.). **Starting a Statistical Quality Control Program at Swift & Company, pp. 221–31.** In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1854. **WILCOXON, FRANK** (Group Leader, Insecticide and Fungicide Laboratory, American Cyanamid Co., Stamford, Conn.). **Probability Tables for Individual Comparisons by Ranking Methods.** *Biometrics* 3(3):119–22 S '47.*

1855. **WILKS, S. S.** (Professor of Mathematics and Director of the Section of Mathematical Statistics, Princeton University, Princeton, N. J.). **Crisis in Statistical Personnel: Summary of a National Research Council Report.** *Am Statistician* 1(2):8–9+ O '47.*

1856. **WILKS, S. S. Order Statistics.** *B Am Math Soc* 54(1):6–50 Ja '48.*

1857. **WILKS, S. S. Personnel and Training Problems in Statistics.** *Am Math Mo* 54(9):525–8 N '47.*

1858. **WILKS, S. S. Statistical Training for Industry.** *Ind & Eng Chem, Anal Ed* 19(12):953–5 D '47.*

1859. **WILLIAMS, C. B.** (Head, Department of Entomology, Rothamsted Experimental Station, Harpenden, England). **Recent Light Trap Catches of Lepidoptera in U.S.A. Analysed in Relation to the Logarithmic Series and the Index of Diversity.** *Ann Entomol Soc Am* 38(3):357–64 S '45.*

1860. ★**WILLIAMS, JEANNE FREEMAN** (Instructor in Mathematics, Fort Trumbull Branch, University of Connecticut, New London, Conn.). **Methods of Using Meteorological and Soils Data in Experiments Involving Certain Vitamins and Minerals.** Master's thesis, North Carolina State College (Raleigh, N. C.), 1946.

1861. **WILSON, R. H.** (General Service Manager, Bell Telephone Laboratories, Inc., 463 West St., New York 14, N. Y.). **Scientific Control of Stockroom Inventories, pp. 175–92.** In *Conference Papers: First Annual Convention of American Society for Quality Control . . .* (*see* 1380).

1862. **WINSOR, CHARLES P.** (Assistant Professor of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.). **Quotations: "Das Gesetz der Kleinen Zahlen."** English text. *Human Biol* 19(3):154–61 S '47.*

1863. **WINTNER, AUREL** (Professor of Mathematics, Johns Hopkins University, Baltimore, Md.). **On the Shape of the Angular Case of Cauchy's Distribution Curves.** *Ann Math Stat* 18(4):589–93 D '47.*

1864. **WISHART, JOHN** (Reader in Statistics, University of Cambridge, Cambridge, England). **Note on the Probability Distribution Arising in the Study of the Institute's Examinations.** *J Inst Actuaries Students' Soc* 6(3):140–3 Ja '47.*

1865. **WISHART, JOHN. Proof of the Distributions of χ^2, of the Estimate of Variance, and the Variance Ratio.** *J Inst Actuaries Students' Soc* 7(2):98–103 O '47.*

1866. **WISHART, JOHN. The Variance Ratio Test in Statistics.** *J Inst Actuaries Students' Soc* 6(4):172–84 Ap '47.*

1867. **WOLD, HERMAN O. A.** (Director, Institute of Statistics, University of Uppsala, Uppsala, Sweden). **Estimation of Economic Relationships.** Abstract. *Econometrica* 16(1):33–6 Ja '48.*

1868. **WOOD, ERIC C.** (Chief Chemist, Virol Ltd., Hangar Lane, Ealing, London W.5, England). **Short Cuts to the Estimation of Standard Errors, Particularly in Microbiological Assays.** *Chem & Ind* (24):334–6 Je 14 '47.*

1869. **WOOD, E. C. Some Statistical Considerations in Organoleptic Tests.** Abstract. *Chem & Ind* (7):107–8 F 14 '48.*

1870. **WOOD, WILLIAM V.** (President, London Midland and Scottish Railway Co., Euston, London N.W.1, England). **The Value of the Actuary in General Statistical and Administrative Work.** *J Inst Actuaries* 73(336):109–12 Je '47.*

1871. WOOFTER, T. J. (Director of Research, Federal Security Agency, Washington, D. C.). Completed Generation Reproduction Rates. *Human Biol* 19(3):133–53 S '47.*

1872. YATES, FRANK (Head, Statistical Department, Rothamsted Experimental Station, Harpenden, Herts, England). The Influence of Agricultural Research Statistics on the Development of Sampling Theory. Abstract. *Econometrica* 16(1):69–71 Ja '48.*

1873. YERUSHALMY, JACOB (Professor of Biostatistics, School of Public Health, University of California, Berkeley 4, Calif.). Statistical Problems in Assessing Methods of Medical Diagnosis With Special Reference to X-Ray Techniques. *Pub Health Rep* 62(40):1432–49 O 3 '47.* For discussion, *see* 1711.

1874. YOUDEN, W. J. (Physical Chemist, Boyce Thompson Institute for Plant Research, Inc., Yonkers, N. Y.). Technique for Testing the Accuracy of Analytical Data. *Ind & Eng Chem, Anal Ed* 19(12):946–50 D '47.*

1875. YULE, G. UDNY (Formerly Reader in Statistics, University of Cambridge, Cambridge, England). On a Method of Studying Time-Series Based on Their Internal Correlations. *J Royal Stat Soc* 108(1–2):208–25 '45.* Note on Mr. Yule's Paper by M. G. KENDALL. *J Royal Stat Soc* 108(1–2):226–30 '45.*

1876. ZEISEL, HANS (Associate Director of Research, McCann-Erickson, Inc., 50 Rockefeller Plaza, New York 20, N. Y.). Coordinating the Measurements of Radio Listening. *J Am Stat Assn* 42(240):512–22 D '47.*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 43 JUNE 1948 NUMBER 242

ARTICLES

The War Production Board's Statistical Reporting Experience . . . . David Novick and George A. Steiner 201

On Pooling Data . . . . Frederick Mosteller 231

Correction
Use of Variance Components in the Analyses of Hog Prices in Two Markets . . . . R. L. Anderson 242

On Estimating Precision of Measuring Instruments and Product Variability . . . . Frank E. Grubbs 243

Statistics in Insurance
I. Life Company Investments and the Capital Markets . . . . Fergus J. McDiarmid 265
II. Mortality Statistics for Life Insurance Underwriting . . . . Edward A. Lew 274
III. Statistical Problems in Developing Voluntary Medical Expense Insurance Plans . . . . Morton D. Miller 290

Statistical Methodology Index, No. 12 . . . Oscar Krisen Buros 356

PROCEEDINGS OF THE 107TH ANNUAL MEETING

Program 299
List of Committees and Representatives for 1947 307
Minutes of the Annual Business Meeting 308
Report of the Board of Directors 311
Report of the Secretary 314
Report of the Treasurer 314
Report of the Auditors 315

BOOK REVIEWS

Davies, Owen L., editor. *Statistical Methods in Research and Production: With Special Reference to the Chemical Industry* . . . . D. B. DeLury and A. E. R. Westman 318
. . . . E. L. Lehmann 320

Eisenhart, Churchill, Millard W. Hastay, and W. Allen Wallis, editors. Statistical Research Group, Columbia University, *Selected Techniques of Statistical Analysis: For Scientific and Industrial Research and Production and Management Engineering* . . Henry Scheffé 321

Fisher, Ronald A., *The Design of Experiments, Fourth Edition* . . . . A. E. Brandt 323

Furlan, L. V., *Das Harmoniegesetz der Statistik: Eine Untersuchung über die metrische Interdependenz der sozialen Erscheinungen* . . . . Hilda Geiringer 325

Kelley, Truman Lee, *Fundamentals of Statistics* . . . Cyril Burt 328

. . . . Horace W. Norton 330

Kenney, John F., *Mathematics of Statistics*, Part One, Second Edition . . . . Howard Levene 332

. . . . Ronald W. Shephard 334

Peatman, John Gray, *Descriptive and Sampling Statistics* . . . . Leon Festinger 335

. . . . Frederick F. Stephan 337

Scarborough, J. B., and R. W. Wagner, *Fundamentals of Statistics* . . . . Leo Aroian 339

Smart, L. Edwin and Sam Arnold, *Practical Rules for Graphic Presentation of Business Statistics* . . . . Rudolf Modley 341

Smith, Edward S., *Control Charts: An Introduction to Statistical Quality Control* . . . . Joseph M. Juran 343

. . . . Mason E. Wescott 343

Thionet, Pierre, *Méthods Statistiques Modernes des Administrations Fédérales aux États-Unis* . . . . W. Edwards Deming 346

Westman, A. E. R., *An Extension Course in Statistical Quality Control* . . . . A. C. Cohen, Jr. 350

. . . . Paul Peach 351

Zetterlund, Stig *Toleranser och Passninger*. . . . Cecil C. Craig 352

Conference Papers: *First Annual Convention, American Society for Quality Control and Second Midwest Quality Control Conference, June 5–6, 1947* . . . . H. A. Freeman 353

. . . . Eugene L. Grant 354

Index to Volumes 1–34, 1888–1939, may be obtained from the ASA. The Journal is also indexed in the Industrial Arts Index, the International Index to Periodicals, and the Public Affairs Information Service Bulletin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 242 JUNE 1948 *Volume 43*

THE WAR PRODUCTION BOARD'S STATISTICAL REPORTING EXPERIENCE

DAVID NOVICK,*
University of Puerto Rico
AND
GEORGE A. STEINER,†
University of Illinois

This is the first of a series of four articles that will present an exploratory analysis of the experience of the War Production Board and its predecessor agencies in developing and using data from industry for policy and administrative purposes during World War II. The present article describes the growth of the WPB's reporting structure.

INTRODUCTION

INTELLIGENT wartime industrial control is impossible without adequate methods for collecting and using facts. In establishing policy for control over industrial production and distribution and in administering these controls, it is necessary for government continuously to make choices among competing demands for resources at all levels of economic activity. Facts and factual collection methods are managerial tools without which intelligent choices among alternative courses of action cannot be made.

When the occasion arose for the national government to direct the industrial machine of this nation in World War II, there existed neither the data necessary to support basic policy decisions and administrative action nor the technical statistical discipline by means of which the data could be produced. The task of industrial mobilization and continuous control of the American industrial machine during the war was unbelievably complex and difficult. But it was a problem immeasurably complicated by the absence of an efficient emergency industrial statistical reporting system.

---

* Formerly Director, Program and Statistics Bureau, War Production Board.

† Formerly Director, Industrial Statistics Division, Civilian Production Administration.

In 1940 and 1941 it became clearer daily to those administering the industrial production program that neither the necessary facts nor satisfactory methods for getting them were available. Peacetime industrial statistical series were, for purposes of wartime mobilization, incomplete in coverage, not timely, deficient in penetrating detail, subject to serious inaccuracies, and uncoordinated. In addition, peacetime methods of collecting and analyzing industrial series in many instances were found inadaptable to the needs of war.

These deficiencies, although lamentable, are understandable. Peacetime factual needs, and methods and series to support them, are far different from those required for wartime administration. In time of peace the economy of this nation is not a centrally directed one but rather one in which the totality of economic activity is based upon a multitude of individual transactions free of government restrictions except those of a general policing character. As a consequence, no need exists for the government to collect data which reflect detailed economic actions. Rather, government is concerned more with information about over-all economic activity and data necessary to carry on specific administrative programs of various regulatory commissions and boards.

This type of data and many of the methods used to collect them are not those required by war. In periods of crisis, national policy demands that the sum total of business operations accomplish specific and well-defined objectives. To meet these objectives the government is forced into day-to-day relations with business units. Government administrators, as a result, need to collect data showing a balanced factual picture of the industrial machine, in which each detail is related to the total and to every other. Business management requires a continuing flow of accurate and detailed information to make the kinds of decisions needed to conform to government regulations and to effect the maximum contribution of the concern's resources to the national objectives. These are crisis needs.

Prior to 1940 the Federal government had two experiences with factual needs of crisis administration. As part of the World War I program of industrial control, the Federal government found it necessary to introduce a large variety and number of industrial data reporting systems. These provided information necessary for the formulation of policy decisions and administrative actions relating to the distribution and production of specific materials and products. At the end of the war practically all of the wartime reporting structure was abandoned, and many of the lessons were buried with the dead files of that conflict.

When the national economic machinery broke down in the 1930's, the government found it necessary to expand its concept of public welfare to include both control over and direction of private business so that the net effect of business operations would be centered upon economic recovery. As an inevitable part of the regulations of such emergency agencies as the National Recovery Administration and the Agricultural Adjustment Administration, the government again developed a complex industrial statistical reporting system by means of which it could support basic policy decisions and administrative actions relating to business operations. With the invalidation of the NRA and the AAA, the data collections were not continued by either government or industry and no study of emergency statistical methods was undertaken.

As a result, there was not available at the beginning of World War II the industrial statistical reporting discipline that emergency administration was found to require. Starting from this point, industrial mobilization administrators began through experimentation, improvisation, and trial-by-error methods the laborious and costly task of constructing an industrial statistical reporting framework suitable to meet emergency management needs. It was only through bitter and time-consuming struggle that the War Production Board eventually developed methods and techniques by which its comprehensive reporting requirements were reasonably adequate to meet its most urgent policy and administrative needs.

In a matter of days following VE-Day the great majority of the WPB reporting instruments were discontinued, as they should have been. In peacetime years the government avoids day-to-day operating relationships with business. This also is as it should be. As a result, both the wartime statistical series and the methods and techniques developed to acquire and use them will tend to disappear. Undoubtedly some of the wartime experience will be incorporated into industrial and governmental peacetime statistical collections and methods. But the integrated discipline of a satisfactory wartime industrial statistical reporting system will find no peacetime use.

This inevitable peacetime disuse must not result in unpreparedness for a possible future emergency. Adequate preparation must embody preservation of the principal lessons of World War II concerning factual tools and procedures required to manage successfully the economic system. But merely preserving the lessons of World War II experience is not enough. Many important tools of administration were still crude and blunt at the end of the war despite the whetstone grinding of five long war years. In time of peace these tools must be studied and improved so they will be razor-sharp for immediate use in any emergency.

Study of the experiences of the WPB in collecting, analyzing, and using industrial statistics, if executed at the proper level of professional proficiency, can save the nation both millions of dollars and years of time in improving its peacetime industrial reporting systems. Perhaps of more importance, neglect to take advantage of and build upon this experience will more than likely leave the nation, on the impact of any future emergency, without the factual tools required to administer economic matters. Failure to have available for immediate use in a future crisis an adequate reporting system may be catastrophic if the military movements in that emergency involve the sudden devastation now attributed to atomic bombs, supersonic projectiles, and guided missiles.

The authors have tried to avoid the use of footnotes. The great bulk of source material, aside from individual questionnaires, is located in personal memorandum of the authors and memorandum of various WPB officials. Both sets of documents are generally unavailable. For those interested in pursuing further the analysis of methodology evidenced in individual questionnaires, attention is drawn to the *Catalog of War Production Board Reporting and Application Forms,* prepared by the Industrial Statistics Division of the Civilian Production Administration in 1946. This *Catalog,* published in ten volumes by the U. S. Government Printing Office, reproduces and analyzes approximately 1,200 of the WPB's most important questionnaires.

PART I

GROWTH OF THE WPB REPORTING STRUCTURE

To satisfy the needs both of policy-makers and of administrators for information to control production and distribution, the WPB created and used a reporting structure of unprecedented scope and complexity. This reporting structure was not built. It was assembled. It grew out of the needs of the moment; it was altered to reflect rapidly shifting objectives, organization, policy, and personnel; and it was molded from the compromise of powerful opposing influences. The problems encountered in its development were as complex and broad as the industrial production system it was designed to harness. No single discipline or procedure ever animated all of its supposedly systematic parts. Its techniques were never completely mastered and its potentialities never fully realized. At no time did it fully satisfy the fundamental needs of policy and administration.

The purpose of this section of the series is to reduce this vast and intricate agglomeration of fact-gathering devices to understandable proportions and by so doing to establish a background which may serve for a presentation of selected fundamental technical problems and lessons connected with WPB statistical methods. To accomplish this, there is presented a broad panoramic sketch of the principal movements taken by the reporting structure through the war period. It is an analysis of the struggle to develop a reporting system to produce facts for wartime management, the principal problems faced in the struggle, and the attributes of the methods introduced to reach the goals set.

Closing Informational Gaps in 1940. In the year 1940 the day seemed remote when the United States would have to reconstitute its industrial system into a single-purpose war-production machine. Defense production for our Allies and our own armed forces at first was comfortably accomplished within available productive resources. Only gradually and in individual areas did it encounter resistance. As a consequence the importance of an adequate statistical discipline to support industrial mobilization management became apparent rather slowly. Actions taken to develop a sound statistical order were at first, therefore, essentially in the nature of stop-gap measures.

The first comprehensive factual research by the National Defense Advisory Commission was designed to establish broad industrial mobilization policy guides. Efforts were made to develop facts about the magnitude of economic resources available for war and nonwar demands, the level at which specific resources might cease to be adequate to satisfy potential demand, and the industrial and geographical areas in which new plant facilities should be built and located. Existing data proved to be inadequate for this work even though completion of the studies required only the broadest type of information.

Failure of this approach led to another which turned out to be even less fruitful. Military agencies were urged to develop complete statistical pictures of their procurement and the claims which they were currently making and contemplated placing in the future on scarce materials and products. This, together with a comparable statement of nonmilitary demand, it was correctly reasoned, could provide much of the basic information needed for broad industrial mobilization policy decisions. By the time useful action was taken in developing such overall statements of claims on industry, the United States was already engaged in war!

The earliest demand for detailed industrial data centered upon plant operations and capacity. Recognizing the paucity of figures in government about such matters a twofold effort was made to get facts from industry. First, various State Industrial Councils obtained information on plant capacities and equipment within their states. Second, the Bureau of Labor Statistics was asked to develop a report on individual plant operations by which the NDAC could learn the extent to which plants were actually operating 24 hours a day, seven days a week. The State Councils accumulated a series of reports and submitted them to the NDAC but the results were inadequate for solving the statistical problems of that agency. The Bureau of Labor Statistics developed a questionnaire which requested monthly reporting from about 300 plants in the metal-working trades concerning man-hours worked, labor shifts, and weekend employment. Both programs fell far short of the factual requirements of the times.

Hindsight reveals quite clearly that such amateurish methods could not provide the information required for a war. There was, however, no dearth of suggestions for getting the information needed. In the spring of 1940 and continuing into the autumn, for example, there was considerable discussion in the NDAC of the desirability of collecting from industry highly detailed reports as a part of what was vaguely referred to as "a modified German Plan." It was assumed that the Germans, who were much deeper in war than the United States, had developed a machine by machine, man by man, pound of material by pound of material, weekly or daily reporting system which by some type of tabulating magic was quickly converted into both an all-out summary for policy-making and a completely detailed report for administrative decisions. The apparent success of German industrial mobilization led to the idea that the United States should have the same sort of statistical system. Introduction of this type of reporting system, however, inevitably would have been labeled a "bureaucratic deluge of paperwork" at this stage of the war effort. For this and other reasons the idea was abandoned.

Individual Preference Ratings. The search for information could not be allowed to delay the solution of current problems concerning acquisition of scarce materials needed to produce vital defense equipment. To meet these issues, the Army and Navy Munitions Board, with but limited legal and executive authority, established in August 1940 a priority plan based on preference-rating certificates. By means of these documents (PD-1, 2, 3, 4 and 5) it became possible for producers holding government contracts for defense equipment to apply for, receive, and extend preferred treatment to items they and their suppliers needed to fulfill their contracts. A *Priorities Critical List* was established to limit the materials and products to which preference could be applied. The listing was relatively short and specified items in broad, general terms, such as: barometers, bombs, cranes, explosives, chromium alloy steel, aluminum and aluminum alloys, and tin.

With a broadening and acceleration of defense production and expansion in the materials and products for which preference was applicable, this system operated less effectively with each passing day. So long as sufficient materials and productive capacity existed to support both defense and unrestricted civilian production the system was effective in assuring to manufacturers with government contracts the materials they needed for production. But, as defense production mounted, it became increasingly difficult to control preference ratings so as to hold preference-rated demand for critical materials within available supply. The inevitable result was the rise of log jams among competing authorized demands in more and more products and materials.

These log jams could not be broken by the individual preference-rating system. The system contained no automatic adjudication of competing claims at any level of manufacture when supplies of materials and products were insufficient to fill all rated demands. In addition, the rating certificates could not be used to produce the facts government could use to resolve the problems by administrative action. Not only was there little uniformity in data incorporated in rating certificates, but the flow of paper reached such proportions that it was impossible to use the data for comprehensive policy or broad administrative action.

Breaking the Log Jams. The first efforts to break these log jams followed admonitory rather than mandatory patterns. Policy decisions made on the basis of whatever data were available were translated into detailed regulations and control orders for which there was no implementation in the form of reports transmitted between government and industry. It was only a matter of months, however, before problems in connection with such tight items as nickel-bearing steel, magnesium, aluminum, machine tools, and heavy industrial equipment became so acute that the inadequacy of generalized self-administered controls to effect the kind of production and distribution control required by the situation was generally recognized. Experience demonstrated that only the most detailed type of administration and centralized control could be expected to provide the desired results.

In the autumn of 1940, the need for new detailed information to solve critical problems connected with individual products and materials had become so urgent that a fresh approach to the acquisition of data was of paramount importance. Each of the Divisions of the

NDAC,[1] and later the Office of Production Management, charged with responsibility for control of a particular material or product worked out new and more detailed programs for the collection of data required for daily administration of production and distribution controls. Those responsible for magnesium, for example, sought data about all the direct and indirect economic problems involved in the production and use of magnesium. The same pattern was followed by those responsible for nickel, alloy steel, machine tools, and other items in short supply. Centralized control over questionnaire issuance fortunately had been introduced in the early days of the NDAC and continued into the Office of Production Management. As a consequence, surveys actually introduced as a result of these programs were limited to those most obviously needed to solve an immediate problem. The country was thus spared the painful experience of innumerable data requests which inevitably would have overlapped, duplicated, and in other ways interfered with one another, not only in the data collection itself but also in the analysis and administration of the results. A trend was begun, however, to acquire more and more data about individual products and materials, each with little reference to the other.

In March 1941 a more realistic approach was made by the recently formed OPM (Office of Production Management). This involved modification of the preference-rating system by the introduction of P, or "blanket preference rating," orders, providing preference-rating assistance to classes of products; and M and E orders, directing the distribution of basic raw materials and equipment. In September 1941 L orders were introduced to facilitate operation of M orders by limiting non-essential demand for materials under M-order control. The great bulk of information pertaining to individual products and materials which was acquired throughout 1941 accompanied administration of these orders.

P Orders. P orders provided for the assignment of ratings to classes of products, usually at the final product end of the industrial system. One type of P order permitted automatic assignment of ratings for procurement of material needed in the production of the classes of end products specified in the order. Such were the first orders, namely, P-1 applicable to electric traveling cranes, P-2 applying to machine tools, P-3 for airframes, and P-4 for airplane engines and propellers. Each order, with few exceptions, specified a reporting form upon which producers to whom orders were issued had to report periodically to the OPM. In addition, producers using or receiving preference treatment under an order were required to maintain records as specified in the order. This provision, however, was often vague and carried no compulsion of compliance. This type of P order, on the whole, was more succesful in relieving a tottering preference-rating system than in developing satisfactory new quantitative measuring tools and methods for using them.

---

[1] Throughout the text all subdivisions of various agencies are specified as Divisions, for purposes of clarity, except in a few instances in which quotations from orders incorporate the official title of a subdivision which is other than a division.

The Defense Supplies Rating Plan, issued under order P-6 in May 1941 was a different type of order and its methods were destined to become of vast importance in WPB operations. Certain manufacturers, such as those engaged in producing electrical motors, twist drills, and mill supplies, received defense orders indirectly and on shorter delivery cycles than their manufacturing time. Deliveries customarily were made from stock, and replenishment of scarce materials for production of additional stock could not be made by extensions of preference certificates for the small amounts of separate materials going into each individual motor or twist drill. The application of individual preference ratings to these producers was quite unrelated to the manufacturing procedures of their products. It was the purpose of P-6 to facilitate production for such manufacturers.

Each manufacturer applying for aid under the order presented to the OPM, on form PD-25, the following information for his plant or subdivision thereof; past quarterly material consumption, material stocks on hand, and current quarterly requirements for those scarce materials for which assistance in acquisition was requested; the amount of defense business; and the specific finished products for which scarce materials were needed. This plant questionnaire embodied rudimentary methods from which profitable development was to be made.

Principles originated by the Defense Supplies Rating Plan were used again in August to aid manufacturers of various health supplies. Order P-29, Health Supplies Rating Plan, was issued to aid manufacturers of specified health equipment who were encountering difficulty in obtaining ratings sufficiently high to procure basic materials needed in their manufacture. Form PD-79, used to administer the plan, required from manufacturers plant reports which requested data about finished products to be made from scarce materials, quarterly consumption of and requirements for scarce materials, inventories of scarce materials, and withdrawals from stocks. With these data, as with the information acquired under the Defense Supplies Rating Plan, the OPM was in a position to appraise the aggregate needs of each plant and to issue ratings accordingly.

To broaden the scope of the Defense Supplies Rating Plan, order P-90 was issued in December 1941 to launch the (voluntary) Production Requirements Plan. Its application form, PD-25A, represented a more advanced development of both PD-25 and PD-79. But data requested on PD-25A were also designed to enable a manufacturer to present a complete picture of his operations in relation to the war program and to essential civilian needs, and to state his production-material requirements for these purposes. So successful was the operation of this plan that by early 1942 approximately 7,000 plants were obtaining priority assistance under it.

These plans operated on the basis of an individual manufacturing unit. They recognized the fact that industry was made up of segments of companies and that producers were usually engaged in the manufacture of more than one class of products. The plans, therefore, were designed to provide priority assistance on a plant basis to facilitate procurement for a plant. Methods designed to acquire facts for these purposes were important in paving the way for future comprehensive plant reports.

Rise of Individual Material Allocation Systems. In March 1941 the same month P-1 was issued, the first M order appeared "to conserve the supply and direct the distribution of aluminum." It was followed in succession by M orders for magnesium, ferro-tungsten, tungsten metal powder, neoprene, nickel-bearing steel, and numerous other critical basic materials. (E orders directed distribution of finished products, such as machine tools.) These orders embodied the idea that control of the distribution of basic materials through the industrial system held the key to maximum war production. To administer a system incorporating this idea, producers of basic materials were required to file reports, usually monthly, requesting authorization to ship mill shapes and forms to customers. These reports requested detailed data on material shape and form production, orders by individual customers, and the products for which materials were ultimately to be used. Upon the basis of this information, administrators responsible for materials under control distributed available supplies among competing programs and products.

The dominant characteristic of control policy and industrial reporting in 1941 centered about hundreds of little universes circumscribed by these orders. Each system in use followed a rough pattern but so differed in details from other systems that uniform control and statistical reporting were completely absent.

The steel plate allocation system represents a reporting system designed to control the flow of semi-finished products through a fabricating bottleneck. As eventually developed, the reporting system required each consumer of steel plate to list his requirements with his supplier one month in advance, specifying government contract number, the use of the material, the preference rating, and the order number. A copy of this form (PD-298) was transmitted to the Steel Division, together with a statement of steel plate consumption and inventory (form PD-299). The steel plate producers, in turn, listed and sent to the Steel Division (form PD-169) all orders received and specified those which were included in the rolling schedule for the following month. The Steel Division surveyed these orders and rolling schedules to determine whether any important military requirements had been omitted. In addition, an effort was made to review the schedules to insure the full utilization of the productive capacity of the steel plate mills. Reporting and analysis were centered only upon the needs of the Steel Division to make its decisions. These decisions were made in the absence of complete information about the status of aluminum, copper, and other materials, competing requirements in other areas, component schedules, assembly schedules, etc.

Most individual allocation systems operated on the basis of detailed reports from industry. The typical material report form, or application for authority to ship raw materials, requested all or part of the following type of information from a material producer, on a monthly basis: (a) the part into which the material was to be incorporated; (b) the product into which the part was eventually to be assembled; (c) the use to be made of the product; (d) the government contract number, if any, of the finished product; and (e) other information, such as preference rating and end-use pattern. Producers generally submitted this information on a series of forms. All orders provided for a report, some for more than one. In addition to producer reporting, allocation systems frequently provided for reports from consumers interested in acquiring raw materials under control, warehouses stocking materials, dealers, and others through whom the material passed in its metamorphosis from crude form to final product.

Among all the competing allocation systems there ran the thread of end-use. Most producers' reports requested data on the use which industry expected to make of raw materials, e.g., carbon steel bar to be made into a valve stem for a specific type of valve for a destroyer. Analysis of various forms created for the most important critical materials shows that there were as many different end-use reporting requirements as forms. Each differed from the other. The fact of nonintegration and nonuniformity of end-use requests was negligible beside industry's inability to supply the desired data. Most of the information requested was known only to the assembler of the finished end product, and only a well-established end-use reporting mechanism could transmit the information through the contractual chain. But no such transmission mechanism existed.

Each allocation system, operated autonomously by each Division, resulted in unbalanced flows of materials. Quantities and delivery dates authorized by the Steel Division might not bear any relationship to those authorized by the Copper Division, because neither the mills nor the Divisions consulted one another about the coordinated needs of the purchaser. Nor were they in any position to do so. Each group responsible for a material or product attempted to appraise the needs for its controlled item alone and to direct the use of its supplies through the industrial system in terms of what were considered to be priority needs, with little reference to operations among other allocation systems.

Although M orders provided better control over the distribution of scarce raw materials than preference ratings alone, they did not provide the balance that only a comprehensive over-all reporting system could achieve. Each group responsible for administration of an M order generally was willing to acknowledge the need and advantage of an aggregate statement of demand and supply covering all materials and products. But when this intellectual willingness was pressed with the suggestion that individual controls and specialized allocation systems should be abandoned in favor of aggregate reporting and over-all allocations, the bridge could not be crossed. The proposal was sacrificed to personal desires and individual notions of how control should be exercised. No support for over-all comprehensive reporting and allocations was given by those who were skilled in their specialized fields and who had erected their own control structures. Indeed, to a very large extent any coordinated effort to develop comprehensive controls was quickly torpedoed by the managers of the individual schemes, who saw in aggregate reports a diminution in their individual importance as controllers of segments of the economy. The maintenance of each individual control, they of course felt, was "absolutely essential to maximum production."

These personal feelings cannot be construed as indictments of the patriotism of these men. On the contrary, these administrators were extremely patriotic in much of their desire to cling to the control systems they had established. They knew their materials and they knew the industries through which they moved. They were afraid that persons with less specialized knowledge might jeopardize the national interest not by malice but by ignorance.

Increasing difficulties in hammering down urgent demands for critical materials to match supplies led, in September 1941, to the introduction of L orders. L orders were basically designed to restrict production of certain classes of products to a stated quantity in a given period of time and to limit or prohibit the use of certain specified materials or products in the production of other products. They were preeminently a device by which less essential demand for products and materials under M-order control might be easily eliminated. L orders generally provided for industrial reporting, and as a result were fruitful sources of information about the products and materials to which they applied. Issued in limited numbers in the fall of 1941 (there were 26 issued up to January 1, 1942) they increased rapidly after the end of the year and by June 1942 numbered 154.

Reporting instruments accompanying L orders varied considerably since the specific provisions of different orders pertained to specialized problems of product and material limitations. Some orders required no reporting or applications; others used but one instrument; and some used a number of reporting forms. L-117 which limited production and distribution of heavy power and steam equipment, used (in August 1942) ten different statistical and application forms; L-41, pertaining to construction, utilized eleven reporting instruments. Through such reporting instruments the WPB was able to acquire considerable detailed information about the production, inventories, distribution, and use of products and materials covered by L orders.

Comprehensive Inventory Surveys. The only industry-wide report introduced in 1941 grew out of problems connected with inventory control. Early in the national defense program, manufacturers were impelled by every sound peacetime operating criterion to accumulate inventories of basic materials. By 1941, a point was reached at which a balance had to be struck between increasing defense materials requirements on the one hand and efficient utilization of tightening material supplies on the other. The uncomfortable spread between material supplies and defense requirements for them, already existing in the case of a few basic materials, threatened to expand rapidly to many materials. Facts about use and consumption of inventories, beyond data obtained through operation of M orders, were essential to permit government to flush out idle stocks to help restore balance in the materials demand-supply equation.

Combating this situation demanded facts—facts collected on a comprehensive, industry-wide basis. But again, half measures were tried first. In the beginning, the OPM limped along with spot inventory surveys of a few tight individual materials and the broad data available in the records of permanent government agencies. None of these surveys, however, provided systematic coverage of inventories from primary producers through various processing stages to final fabrication. But the data at hand, piecemeal and fragmentary though they were, showed that inventory accumulation was reaching serious proportions. There was little in them, however, to support specific policy decisions and administrative action to cope with the situation.

The first effort to solve the problem, through General Materials Order No. 1 issued in May 1941, was completely ineffective. Under the provisions of this order, which covered 29 materials, users of the materials specified filed a monthly statement (PD-19A) that they were not increasing inventories, and producers filed statements (PD-19B) that they were not shipping to idle inventories. The deficiencies of this order quickly became apparent, and other measures had to be taken to make operative the ideas embodied in it.

To establish a factual basis upon which General Metals Order No. 1 could be made effective, form OPM-69 (inventory and consumption of scarce materials) was sent in August 1941 to all manufacturers who were holders or users of any of the 70 scarce metals enumerated in the form. A similar form, OPM-75, was tailored to fit the needs of all dealers in and distributors of scarce metals. In addition, comparable information was required on forms OPM-132 and OPM-133 from financial houses, public utilities, and other groups who held these metals. OPM-69 was mailed to about 70,000 manufacturers; the mailing list for the other related forms aggregated approximately 30,000 names. This was the first industry-wide reporting effort in the war period.

Those to whom the forms were sent were asked to report the quantity of each metal which they had used in August 1941 and the amount on hand at the end of that month. Definitions of metals for which data were requested were broad and extended to the point where the raw materials lost their identity as raw materials and became fabricated products. Fabricated product definitions were also vague and inadequate.

The immediate results of OPM-69 and its related forms were rather disappointing, largely because of inexperience in the development of the form itself and problems connected with its tabulation. Tabulated summaries of OPM-69, as a result, were delayed until December 1941. In the meantime events had moved so rapidly that the data reported on OPM-69 and its related forms were considered obsolete by the policy-makers and administrators for whom they had been collected. Not only had difficulties encountered in production moved far beyond the point where inventory accumulation problems were of dominant concern, but the Divisions of OPM had collected inventory and other production information for many individual products and materials on a current basis. Although the OPM-69 data soon lost currency, they contained much information which proved to be useful for a long time to come.

The methods of OPM-69 and its related forms, however, were significant at this stage of the war effort. Of most importance, the issuance of these forms broke the ice, so to speak, in the area of comprehensive reports connected with metal usage and consumption. In addition, much was learned about such technical problems as definitions, units of measure, instructions, mailing lists, and tabulations as they related to industry-wide reporting. Equally as important as the positive methodological lessons were the negative ones which taught what not to do.

Tabulating Preference Rating Certificates. Urgent needs for comprehensive information about many aspects of industrial production centered attention upon the aggregation of data contained in individual preference-rating certificates and resulted in a tabulating effort more ambitious than fruitful. These certificates contained data on materials used in production, manufacturer shipments of defense products, preference ratings, orders, and other sorely needed facts. For a considerable period of time the small volume of these certificates, centrally accumulated in the OPM, was hand-tabulated with some success, but little or nothing was done with those received by the Army and the Navy. As the flow of paper mounted, the Bureau of the Census was asked, in August 1941, to tabulate data by mechanical punch card methods. Thereafter, the flow of paper reached flood proportions and as a result tabulation difficulties grew faster than solutions. Completed summaries were never made and the project was abandoned in January 1942.

The Reporting System at the Time of Actual War. The net effect of all these developments was to leave the United States upon the eve of its active entrance into World War II without the basic statistical-administrative tools needed to cope with pressing production and distribution problems. The core of the most serious problems was demand-supply imbalance for metals. Throughout the fall of 1941 it was becoming more apparent daily that the supply of a number of basic materials, particularly metals, was completely inadequate to satisfy all top priority demands; far too many instances had appeared when even the demands of highest-rated military orders could not be filled from the limited material supplies made available for these demands. The fact was assuming crisis proportions that shortages of vital raw materials, particularly the basic metals, and inequities in the distribution of these materials under the operation of individual allocation and preference rating systems, were throttling the efficient use of the American productive machine. Nothing short of carefully controlled over-all distribution of scarce materials seemed adequate to secure maximum utilization of facilities and full production of essential goods. But this necessitated considerable knowledge about both the demand for and the supply of scarce materials. Neither the tools for acquiring this information, nor those needed to use the information to balance the demand-supply equation existed when this country entered the war.

Notwithstanding the efforts of 1941, available facts about neither side of the demand-supply equation could provide a satisfactory basis for working a balance. Military requirements were unrealistic, inaccurate, and unbalanced within programs; they were more in the realm of "desires" than of "requirements." The OPM, and later the WPB, had progressed little further in developing comparable essential civilian requirements. Civilian programs were disorganized in the aggregate, in detail, and as related to military programs. Available estimates of both were far too high in totality, subject to rapid and unexpected shifts among programs within over-all totals, unadjusted for lead times and contingency allowances, and constantly in the process of changing specification and design. Uncertainties in the measurement of material supplies also created a stumbling block in equating the supply-demand balance. In some cases, calculations of both long-run and short-run estimates of material supply were so poor that later acquisition of accurate figures showed them to have missed by as much as 50 per cent. Overestimates and underestimates were of about equal frequency among various individual calculations. Thus, neither demands nor supplies in aggregate terms were accurate enough for intelligent policy and administrative action.

Inaccuracy, bias, and absence of facts as to both the demand and the supply sides of the raw material equation met head on in early 1942 with completely inept procedures for reconciliation of balance, even had there been sound requirement and supply figures. The WPB had not perfected within its own organization a systematic and effective review of the statistical, technical, and conceptual accuracy of even the poor estimates of demand and supply which were available. There was no administrative machinery for equating demand and supply aggregates at policy levels. There were only the unsynchronized and crude procedures of M orders. Machinery was absent for scheduling raw materials and components through the fabricating chain with some semblance of coordination. As a result, each procurement agency, almost with impunity, heavily mortgaged critical materials, machine tools, labor, and other resources needed to protect its over-optimistic schedules. In the process no reference was made to competing claims of others. The WPB was helpless to reconcile and control these powerful influences, which worked towards imbalances throughout the nation's industrial productive machinery.

After more than two years' experience with wartime production control, facts simply were not available for the most fundamental type of policy decision. This situation was epitomized in the character of policy action taken by the WPB Requirements Committee which, in early 1942, knowingly issued its own over-all allocation directives with full consciousness that no mechanism existed by which it could determine whether its decisions were in full conformance with the over-all industrial situation; by which it could relate its actions to those taken in connection with individual materials or products, either by itself or by product administrators; and by which it could be sure that administration and execution of its actions was adequate for the fulfilment of its decisions. So explosive were pressures at the time that policy decisions had to be made with or without facts on which to base them.

These conditions led to the first concrete attempts made during the war period to obtain comprehensive information by which intelligent over-all metal-allocation policy decisions could be made. Throughout 1941, and continuing into 1942, substantial volumes of information were flowing to administrators of individual products as a result of the operation of P, M, and L orders. To coordinate information about these individual products and materials and to acquire comprehensive information by which war industrial activity could be measured and the demand-supply equation balanced, two industry-wide questionnaires were issued. The first was form WPB-732, which was designed to obtain comprehensive information about operations in the metal-working trades. The second was PD-275, which was designed to acquire plant-by-plant data concerning the use and requirements of critical metals. At the same time renewed pressures were exerted to obtain adequate requirements data from procurement agencies and to accumulate information on essential civilian demands within the WPB itself. These measures stimulated a chain of events which eventually resulted in the accumulation of sufficient data by which policy and administrative actions were reasonably well supported and coordinated for the first time during the war period.

WPB-732. Form WPB-732 traced its ancestry back to the special plant reports collected by the Bureau of Labor Statistics for the NDAC. Upon the entrance of the United States into the war, it was recognized that a central war-control agency could ill afford to be without comprehensive knowledge relating to employment, machine and plant utilization, and other comparable data concerning all important industrial segments. By this time, too, it was clear that the main focus of the WPB control had to be centered upon the metal-fabricating trades. As a result, WPB-732 was inaugurated in March 1942 to obtain such facts from all important manufacturing plants.

WPB-732 was an operations-type statistical report submitted on a comprehensive scale. The original mailing list totaled about 15,000 plants in the metal-working trades but quickly rose to around 25,000. Towards the end of the war the list fell to approximately 9,000 of the largest plants, whose operations constituted about 85 per cent of total dollar shipments in metal fabricating industries. Information was requested concerning the extent to which plants had converted to war work, the extent to which plants were operating at capacity, products manufactured, unfilled orders, employment, machinery utilization, and explanations of failure to keep abreast of war contract delivery schedules.

Submitted on a monthly basis, its primary function was to produce a broad measure of the trends of war production and to reveal, for policy-making purposes, potential areas of production and distribution difficulty. So useful was the information collected from this form and so flexible was its method in adjusting to changing needs of management and to varying practices in industry, that it was destined to outlive the WPB. Throughout its existence, this form underwent constant revision to reflect changes in the preference-rating structure, the introduction of the Production Requirements Plan, and the Controlled Materials Plan; to include new types of data; and to adjust its informational requests to peculiarities of various industries. Throughout all these changes, however, the WPB maintained statistical threads in this questionnaire by which a continuous long-time historical series of comparable and related data could be maintained. This was the only comprehensive reporting instrument of the WPB to embody this feature.

The data collected were at all times valuable guides to policy on, and administration of, the armament production program. As a statistical rather than an application-type report, however, the figures were designed for policy rather than administration and did not constitute a vehicle for direct implementation of policy. But the figures revealed how policy and administrative action were working out in industry; served to detect potential problem areas; measured utilization of manpower and machinery in war work; and provided the basis for hundreds of special statistical reports upon which important decisions were made. Altogether, the data were used to throw light on a wide variety of problems.

PD-275. The introduction of form PD-275 in February 1942 firmly grappled with the materials-control problem facing the WPB. Experiences in late 1941 left no doubt that the prospective drain of an all-out war effort on the country's raw material supplies would become so great that only a coordinated materials-control plan would avoid industrial chaos. It was vital for the WPB to know over-all total requirements for basic materials and the supplies available to satisfy them. These requirements had to be classified according to major uses and time periods, and raw materials had to be scheduled through the industrial system on the basis of some sound order. A major contribution in the acquisition of the facts needed and the methods by which they could be used to exercise control was made by form PD-275.

PD-275 was essentially a modification of PD-25A currently being used under the (voluntary) Production Requirements Plan. Indeed, the form was so prepared that information requested on it was the same as the data asked for under the voluntary PRP. There were two major improvements, however. First, it was sent on a mandatory basis to 11,000 manufacturing establishments whose operations, it was calculated, would require approximately 90 per cent of the metals used in manufacturing. Second, a list of about 110 metal shapes and forms was devised from those under the PRP and preprinted on the form itself. The first change assured coverage; the second assured speed in tabulation. For each of the metals, the form requested data on inventories and consumption in the fourth quarter of 1941 and anticipated requirements in the second quarter of 1942. Manufacturers were asked to submit a separate form for each class of products manufactured if existing inventory and other records were such that separation was possible. This technique facilitated the reporting of data about each product.

The technical results of the survey were heartening. Summary tabulations provided the first reasonably accurate over-all measure of metal consumption and requirements by manufacturing areas. The data revealed large lumps of critical metals lodging in excessive inventories, metal shipments on unrated orders, and metal demands carrying the highest ratings in excess of visible supplies. In other respects the summaries threw into sharp relief the fact that, if urgent war and essential civilian production was to be met, tighter and more efficient control over the flow of metals was necessary.

The WPB, however, was not quite ready for extended mandatory metals control on a comprehensive basis. A second PD-275 was therefore issued in April. This form was sent to 25,000 of the largest metal-using manufacturers and to shipyards, arsenals, railroads, and other consumers of metal. In layout, the form was similar to the previous issue, with one important modification. The form provided space in which manufacturers were required to report metal use and requirements for each class of products. Not only was the coverage of this form better than that of the first but the data obtained were more accurate and more readily tabulated.

The results of the tabulation of this survey reinforced those of the first questionnaire. Both made one fact stand out. The productive capacity of the United States was well beyond its ability to supply critical materials to industry, and, as a result, expansion of the production of the most vital military and civilian goods depended upon the immediate introduction of comprehensive coordinated control over the flow of basic metals. In addition, both forms, together with techniques built up through the operation of the (voluntary) Production Requirements Plan, revealed the administrative and statistical methods by which this could be accomplished.

The Production Requirements Plan. The WPB took the obvious step in the introduction of the (mandatory) Production Requirements Plan on June 10, 1942. Under this plan any plant, or division thereof, whose past or anticipated quarterly use or requirements of critical metals exceeded $5,000 was required to apply for rating assistance. The application form was PD-25A, a modification of the original PD-25A only in its incorporation of various technical lessons provided by PD-275. Essentially these were the preprinted metals stub and the direct relationship of metal use and requirements to individual final products. Exceptions to filing this application form were granted to United States or other government agencies (not including shipyards, arsenals, factories, etc., which were subject to the covering Regulation), transportation companies, public utilities, and so on. In practice, however, data were gathered by other means for these so-called "exempt groups" so that a complete statement of over-all metal use became possible.

Each plant falling within the above requirements was required to submit data, on PD-25A, on the basis of the smallest breakdown of operations which existing raw material inventory records permitted. The calendar quarter was the time unit and for this time period manufacturers reported: (1) the dollar value of past and anticipated shipments for each final product, analyzed by preference ratings; (2) past consumption and future requirements for each of about 200 specified metal shapes and forms, for all operations of the plant; (3) past and anticipated inventory levels for each of the 200 metals; (4) past consumption and future requirements for each of 200 specified metal shapes and forms, for each final product; and (5) past consumption and future requirements for important components. Other information included rated orders, metal requirements for long periods in the future, and analysis of ultimate purchasers of products.

The first PD-25A, covering operations in the second quarter and requirements for the fourth quarter of 1942, was sent to approximately 32,000 plants. Successive quarterly operations extending through the second quarter of 1943 permitted a stabilization of the mailing list at about 30,000 plants or plant divisions. Of this number approximately 15,000 applications accounted for practically all metal consumption in manufacturing.

PD-25A fully covered industrial plant operations as they related to the consumption and use of metals. With this information, competition for basic metals in each plant could be appraised in terms of both mill supplies and the relative importance of end products. The entire structure of input-output relationships in American industry for the first time was silhouetted for study, analysis, and application to the problems of directing the flow of materials through industry. Further, mere addition of figures from each plant revealed reasonably accurate actual metal usage and potential requirements which, when related to calculated supplies and military and civilian programs, provided the means for the balanced allocation of materials in aggregates and in details by industries, plants, and products.

PD-25A approached more closely than any previous effort the sort of comprehensive reporting methodology needed in wartime. Tabulation of data on PD-25A put the Requirements Committee of the WPB in a position to pare total military and civilian requirements to match critical material supplies. When this was done, reduction in authorizations to buy and use metal was made by WPB administrative divisions among various industry groups on a plant-by-plant basis. These decisions were transmitted to manufacturers by authorizations contained on their PD-25A application. PD-25A ,therefore, not only supported the highest policy decisions concerning metal allocations but also functioned as an instrument of administrative action.

Despite the potential usefulness of the PRP as an over-all control system, neither the preference-rating system nor individual allocation devices used by the WPB were eliminated by the adoption of the PRP. The effectiveness of the PRP approach to equate demand-supply imbalances and to produce order in the flow of materials was seriously impaired for these and other reasons noted below. The development of this new technique was accompanied by the adaptation and partial subordination of old methods rather than their complete abandonment. (This tendency toward superimposition of new controls on old ones was prevalent in the WPB's operations and accounts for much of the complexity and duplication which existed in the statistical structure.) In addition, no great progress was made in improving the techniques for calculating basic material supplies, in coming to grips with the problem of military and civilian requirements at the top procurement levels, or in controlling contract issuance.

Materials Accounting Reports. The effort to obtain a comprehensive system to produce facts to control metal flows brought with it a concept of basic material accountability. It was early recognized that accountability for action was an essential part of administrative order: it was one problem to obtain information upon which to determine basic policy and internal administrative actions to carry out policy, and quite another problem to be certain that internal operations and action at the production front were in conformance with policy. The latter assurance could come only in a system where current records of action were kept.

Toward the end of 1941 and early in 1942, efforts were made to control by accounting methods the mortgages placed on scarce metal supplies. They were not too successful. The adoption of the PRP, however, required the establishment of internal record-keeping systems by which current statements of metal authorizations could be reviewed. In no other way would it have been possible to determine whether the WPB administrators were authorizing materials within the limitations established by the Requirements Committee "pie cuts." As a result, comprehensive materials accounting was undertaken for the first time under the PRP.

The concept of materials accounting control growing out of this experience under the PRP constituted one of the major innovations in the reporting system. So inherently necessary, sound, and successful were the general methods developed that control accounting grew to constitute an important part of many future control mechanisms within the WPB, in other wartime agencies, and in industry.

Vertical Versus Horizontal Allocation Reporting. The core of the production and distribution problem lay in the tying of materials needed by suppliers of parts and subassemblies to end products. In the fall of 1942 these problems assumed new importance, as fabricating facility shortages increased. Consideration of this problem brought a shift in thinking about basic over-all allocation control from comprehensive *versus* individual allocation systems (the PRP had settled that conflict) to comprehensive horizontal *versus* vertical allocation techniques. The concept grew more firmly that the horizontal material allotment approach of the PRP had to be discontinued in favor of a vertical material allotment approach. Thus, rather than allot materials directly to manufacturers at each fabrication stage, it was reasoned that metal flows could be tied tightly to end products if end-product producers allotted metals to their suppliers, the suppliers to their sub-suppliers, and so on down the contractual chain to basic material producers.

Probably the most important shortcoming of the PRP as an allocation device (and the one from which most of its other deficiencies stemmed) was that it could not force a balancing of military and essential programs in relationship to limited metal supplies. The WPB through PD-25A was in a position to match total metal requirements, so far as they were known in industry, against total metal supply. Previously the WPB had relinquished its control over the contractual rights of procurement agencies and therefore was not in a position: (1) to force a balancing of military end programs within themselves; (2) to force a balancing of end programs within total material supplies; and (3) to force a tight distribution of metals and components in accordance with balanced programs. The PRP was forced to carry the burden for controlling the flow of metals and components in the absence of programming and scheduling, a burden it was not designed to carry.

In the absence of balanced military and essential civilian programs, the PRP and PD-25A could not effect the desired flow of materials through the industrial system nor could they govern the scheduling of component flows in the volume and time sequence required. When the total supply of tight materials was calculated by the WPB, requirements as submitted by industry on PD-25A could be matched against it. This system produced facts on metal requirements of end programs for which contracts had been let at each level of manufacturing. These requirements could be matched against supply and allocation made accordingly. No mechanism existed, however, for forcing procurement agencies to pare contracts in conformance with this balancing of demand and supply. As a result, metal allocations made to a component producer, for example, often were not in conformance with the changing needs of end programs. In the absence of control over end programs of military agencies, serious difficulties arose with the PRP method of horizontal plant allocations whether such allocations were intelligently made or not.

These deficiencies of the PRP could be resolved, it was felt, by metal distribution on a vertical basis. With such a system the WPB would in effect create a bank balance for quantities of metal for each procurement agency. Those agencies in turn would write "checks" on their balances and distribute them to prime consumers. These proceeds would constitute a bank balance on the books of prime consumers from which they could make authorizations to their suppliers. Suppliers in turn would make extensions to sub-suppliers and so on. The net result would be, it was argued, that "bank withdrawals" could be kept within available balances and the distribution of materials would be tightly tied to and coordinated with production schedules.

After continuous discussion of alternative allocation proposals from early 1942 through the fall of that year, this approach to the problem of metal allocation was chosen. Among the more dominant schemes discussed were: a modified PRP, a "Warrant Plan," the "General Motors Plan," the "Material Allotment Plan," and the "Steel Quota Plan." Out of the exploratory discussions pertaining to these proposals and the critical materials situation there evolved the Controlled Materials Plan, the principal procedure of which was a vertical method of allocation. This plan was to constitute the only comprehensive metal allocation system in operation throughout the remainder of the war.

The Controlled Materials Plan. The Controlled Materials Plan was announced in November 1942, went into effect experimentally in April 1943 and fully in June 1943. The plan was based upon the theory that control over the production and distribution of a few basic materials —steel, copper, and aluminum—which were known as the "controlled materials," was sufficient to regulate a large part of industrial production. (For operational purposes these materials were broken down into 16 shapes and forms.) Control was exercised on the basis of two principles: (1) authorization of only as much production as available supplies of basic metals would permit; and (2) establishing checks upon the use of these materials to maintain authorizations within limited supplies. To execute this control, four basic steps were followed: (1) determination of demands for materials; (2) determination of supplies of materials; (3) matching demand and supply; and (4) feeding material to scheduled production on a vertical-allotment distribution basis. At the apex of the pyramid up which demand-supply information flowed, and down which allotments were channeled, was the WPB Requirements Committee.

To the Requirements Committee, ultimate purchasers of all manufactured products containing controlled materials presented statements of their demands for those materials in terms of programs, physical metal quantities, and time periods. These purchasers, called Claimant Agencies, such as the Army, the Navy, the Maritime Commission, the Office of Lend-Lease Administration, and the Board of Economic Warfare, established so-called end programs designating equipment ultimately purchased for their account. Determination of material needed to fulfill these programs was facilitated by the acquisition from their prime contractors of "prototype" bills of materials (generally submitted on form CMP-1). Prime contractors in turn requested similar information from their suppliers, and so on down the contractual chain to the first fabricator of mill shapes. Thus, extending up through the contractual chain to the WPB Requirements Committee, detailed material requirements flowed. Controlled material supplies in various time periods were calculated by the WPB Controlled Materials Divisions. On the basis of this information, the Requirements Committee, in consultation with the materials Divisions and the Claimant Agencies determined how demand for various agencies and programs could be reduced to match supply. When this was accomplished the Requirements Committee made lump allotments against the supply to Claimant Agencies and their principal programs.

Upon receipt of allotments from the WPB, the Claimant Agencies distributed quantities among their own detailed programs. Authorizations for all detailed programs had to remain within total quantities received. Each agency, therefore, instead of competing with others through the issuance of preference ratings, worked toward a bank balance. If it wasted materials in one direction, it had less in another. Claimant Agencies then divided their allotments among their prime contractors, giving each enough controlled materials to carry on his part of the total program. Prime contractors, within the limits of their own "bank balances," passed along the portion of their allotments needed by sub-contractors and sub-subcontractors to produce parts and subassemblies for their contracts. Thus, material distribution was passed down through the contractual chain. This simple procedure of following contractual channels tied metal demand to production schedules even though it did not reduce the dollar value of outstanding contracts.

The entire plan did not operate on vertical-allotment distribution. Components and shelf-items such as nuts, bolts, ball bearings, and electric motors were purchased by so many manufacturers for so many end uses that producers could not conveniently obtain controlled material allotments on the basis of vertical-allotment distribution. For these and other manufacturers of products such as pumps, gasoline engines and such wide-use components, who would have had trouble in receiving allotments by vertical distribution, the PRP method of horizontal allocation was adopted into the CMP. By this means, as under the PRP, these producers asked for and received allotments directly from the WPB.

To facilitate operation of the CMP, all finished products of manufacturers were classified into two groups, Class A and Class B. Class B products originally were intended to include only those subassemblies or miscellaneous or specialized items of equipment for which individual allotments could not practically be made from prime consumers down the contracting chain. This concept was modified, however, and many final products ultimately were defined as Class B products. All other products were Class A products.

The distinction between these two classifications was vital in CMP operations. All producers of Class A products were included in the vertical chain of allotment distribution, submitting their requirements (form CMP-4A) to their customers and receiving from their customers, in turn, metal allotments. Class B producers, on the other hand, were included in the horizontal-allotment distribution system, submitting directly to the WPB (form CMP-4B) their requirements for metals and, in turn, receiving from the WPB their needed allotments. As distinguished from the plant approach of the PRP, all CMP application and authorization documents were in terms of individual Class A or Class B products.

Insurance that demands were kept within metal supplies, once the WPB Requirements Committee had distributed allotments, was accomplished by an effective materials accounting system. Establishment of adequate accounting controls throughout the allotment system was based on the idea, born of experience, that effective control necessitated procedures by which any overdrawing of accounts at any point in the distribution process could be prevented or detected. Application of this idea resulted in accounts of record showing for each program, for each material shape and form, for each quarterly period: authorizations, issuances, and balances. Such records were maintained in the WPB Industry Divisions, Claimant Agencies, and every industrial plant receiving allotments.

An effective check on the impact of metal authorizations at the basic-controlled-material-producer-level was accomplished by a system of periodic reports made by producers to the WPB Controlled Materials Divisions. On form CMP-8, producers reported shipments and unfilled orders. These reports were tabulated by the WPB and provided a direct measure of procurement pursuant to allotment by Claimant Agency and Industry Divisions, in terms of end programs for which original allotments were made by the WPB Requirements Committee.

The flow of end-use program information from Claimant Agencies through industry to metal mills was made on the basis of the passage of allotment program symbols. Originally, identification of allotments called for the passage of thirteen digits but the unworkability of the transmission of so many numbers resulted in a modification adopted before the Plan became operative, which provided only for identification of the Claimant Agency and the first digit of its program number.

Benefiting from the experiences of the PRP, various control systems, in effect in the WPB at the time the CMP was introduced were modified to conform with the latter. First, and of paramount importance, acquisition of controlled materials was only on the basis of direct allotment; the preference-rating system did not apply. Second, substantial modification or revocation was made among M, E, and L orders to avoid any interference or conflict with the operation of the CMP. As a result, some questionnaires established to operate individual allocation systems and other programs were abolished or substantially modified. The reporting and application forms of the CMP were given a dominant position which never had been extended to the PRP and its instruments.

At this stage of the war effort, inventories had ceased to be a serious problem. Unbalance between the demand for and supply of critical materials had flushed out most idle usable inventories. The CMP provided no automatic cleansing of idle inventories but rather exerted control by establishing a 60-day limitation on inventory accumulation. To assure compliance a plant report (form CMP-7) was issued covering operations in the third quarter of 1943. The form requested information concerning plant shipments by products and for each CMP metal shape and form: actual use, receipts, and inventory. After one quarter of operation, the basic features of the form were incorporated into form WPB-732.

The CMP proved to be eminently successful in equating the demand-supply formula for controlled materials. Indeed the CMP, perhaps more than anything else, was a fruition of the concept that no WPB control policy over metals was adequate which did not devolve upon procurement agencies the responsibility for balancing their end programs and establishing methods by which these end programs could be translated into elemental common denominators of materials and processing time schedules. The CMP did force such a balancing and capitalized upon it by establishing suitable procedures in the WPB for acquiring accurate supply data and for equating demand and supply along realistic patterns. The CMP unquestionably firmly settled the problem of equating the demands for and the supplies of metals.

The plan was less successful in scheduling production. In operation, the CMP was primarily a materials-control system. It did not provide the machinery for balancing component production schedules in conformance with end programs. Claimant Agencies had the authority to spread out material allotments in conformance with their end-program requirements, but the flow of materials was not controlled in terms of time and volume through the industrial chain. As a result, component production for end programs was often far out of line with real requirements. Efforts to solve this problem were embodied in various regulations designed to supplement the operation of the CMP.

Scheduling Production. There emerged in full force in 1942 a group of problems relating to production scheduling. Although their emergence was detected early in the year, measures for their solution were not to be adopted until 1942 had almost ended. An analysis of four months of ordnance production, for example, indicated that this program was getting out of balance. The output of end items and the output of the complementary items essential to the use of the end items were not proceeding at the same rate. The scheduled production of components for artillery and aircraft ammunition was out of joint with the requirements for these components to meet production objectives for complete items. Many component schedules ran to two or three times requirements, and the schedules for two-thirds of the components called for production more than 15 per cent in excess of requirements. The result was that the ammunition program was using excessive quantities of scarce materials. This and comparable situations demanded administrative processes by which the flow of critical materials would be controlled in accordance with scheduled uses, which should be limited to requirements for production objectives.

As the year 1942 moved on it became increasingly evident that, although total production was advancing, its rate of progress was leveling off. Solution to the problem lay not so much in basic material supplies, which had increased considerably during the year and had been matched with demand under the PRP, as in a system which defined points where materials were needed to effect a balance among end programs and which would direct critical materials to them.

The military services, which had undertaken the scheduling of production for military items, maintained little balance in schedules. Some schedules were over-optimistic, and others were conservative; there was little agreement as to lead factors, conversion periods, and provisions for the effect of changes in design and specifications. Moreover—and this was of crucial importance—there was no central review of the schedules for the purpose of reconciling competitive demands for common resources. And yet without such a review of schedules to align competing demands in relation to the maximum supply of common resources it was impossible to insure the development of realistic schedules. With the development of each new bottleneck in components, the record showed more clearly that inept scheduling had to be rectified if the nation was to squeeze the maximum production out of its resources.

In December 1942 a Production Executive Committee was established in the WPB with broad and inclusive powers to schedule components of all Claimant Agencies; to see that program schedules did not conflict, were in balance, and were consistent with the maximum productive possibilities of the national economy; and to ensure that schedules were in accord with the strategic requirements of the Joint Chiefs of Staff. A series of preliminary steps were taken to solve scheduling and related problems and in February 1943 General Scheduling Order M-293 was announced, to become operative on April 1, 1943.

The original M-293 applied to 35 classes of components classified into three categories of urgency. The WPB Industry Divisions were given full authority to control the shipment of these items. Information to exercise the necessary control was reported by manufacturers. Form PD-900 provided data on production capacity, orders, and order schedules for various periods. Forms PD-901 and PD-902 requested detailed delivery schedules. After the issuance of M-293, additional

WPB Industry Division scheduling orders appeared like dandelions in the spring and brought many new questionnaires with them. Since all these forms required details concerning scheduling of individual orders for individual items and customers, they more frequently than not were bulky and exceedingly burdensome to fill. Most of these orders were popular in industry, however, because by reason of them a manufacturer could freeze shipment schedules and go about his business without the need for constant adjustment and readjustment in schedules otherwise necessitated by the priorities system.

M-293 probably stimulated the flow of a larger volume of reports to the PWB than any other single regulation. After the novelty of component scheduling had worn off, most of the scheduling forms received by the WPB were filed away without review or any action whatsoever.

Reporting Trends in Other Areas. The preceding movements, viewed from an historical perspective, appear to be those of outstanding importance in the evolution of the reporting structure. Apart from them, however, there was constant evolutionary development, of a less spectacular character and with less fundamental impact upon industrial control, in hundreds of individual products and materials. In each instance in which the demand-supply equation for a particular product was so unbalanced as to impede production, special reporting requirements were instituted to provide facts for intelligent administration. When the problem ceased to exist, reporting was generally discontinued. The often unpredictable impact of war upon the thousands of products needed for its prosecution created a substantial ebb and flow of control problems which in turn was reflected in the rise an discontinuance of thousands of individual questionnaires.

The use of control procedures and reporting methods adopted to support them often paralleled and leaned heavily upon the experiences described above. Thus, for example, the methods of the Lumber Control Plan and the Tire Allotment Plan were essentially those developed along PRP lines.

By the middle of 1944 the basic WPB reporting structure was created. Thereafter, its movements were essentially caused by modifications of existing reports rather than by the introduction of revolutionary new methods. With the approach of VE-Day, and later VJ-Day, strenuous efforts were made to simplify the reporting structure by eliminating from it all questionnaires of little or doubtful use. The result was that in the period immediately preceding VJ-Day the WPB operated with fewer questionnaires than in any previous period after the entrance of the United States into active warfare.

ON POOLING DATA

Frederick Mosteller
Harvard University

On the basis of samples from normally distributed variates X_1 and X_2, how should one estimate the mean of X_1? Should one pool or not pool the means on the basis of a significance test, or should one pool continuously, or not at all? It turns out that if the difference between the true means can be thought of as normally distributed from sample to sample, pooling with unequal weights is preferable.

INTRODUCTION

The purpose of this article is to discuss a very simple problem concerning the pooling of data. If we have a sample of n measurements from population A and a sample of n measurements from population B, where A and B are similar, how shall we estimate the mean of population A? Shall we merely take the mean of the sample from A, or shall we attempt to combine the means of the two samples in the hope of improving our estimate of A in some average sense?

A variation of this problem is frequently met in the analysis of variance where various estimates of error are sometimes pooled after an initial significance test has been made. This case has been treated in some detail by T. A. Bancroft. The problem of pooling means does not seem to have been treated. The principal reason for discussing the pooling of means is that the notions are much more easily understood than the corresponding question of pooling sums of squares. Considerable insight into the general pooling problem can be gained by an examination of the more elementary problem. The problem of the pooling of means can be discussed in an elementary class in statistics. In addition, there are occasions when the question of the implications and the results of pooling means arises in practice.

If we have two (or more) similar measures of an individual's performance, is it preferable to take the later measure as an index of his ability, or would it be preferable to average the measures? If a producer sells us two very similar products made of the same materials on the same machines by the same operators, shall we estimate his performance on one of the products by the per cent defective observed on that product, or can we get a better picture by combining the results of samplings on both products? This question may be of considerable interest if the testing happens to be destructive and expensive. These and similar questions are all concerned with the pooling of means. We shall, of course, idealize the problem considerably.

Let us suppose we have two kinds of "12 inch" rulers, which for convenience we designate as "red" and "black." We assume that both kinds of rulers have lengths which are approximately normally and independently distributed with equal known variances σ^2, but unknown means, u_r and u_b, where the subscripts indicate the color of the ruler. We take random samples of n red rulers and n black rulers. Let $\bar{x}_r$, $\bar{x}_b$ be the sample means, then for $k>0$ if

$$|t| = \frac{|\bar{x}_r - \bar{x}_b|}{\sigma}\sqrt{\frac{n}{2}} < k \tag{1}$$

we accept the null hypothesis that $u_r = u_b$, at significance level

$$\alpha = 1 - \frac{1}{\sqrt{2\pi}}\int_{-k}^{k} e^{-\frac{1}{2}t^2}dt.$$

If the inequality in (1) does not hold we reject the null hypothesis in favor of the alternative $u_r \neq u_b$.

Once the significance test has been made, the following question sometimes arises. On the basis of the samples, how should one estimate the mean length of the red rulers?

When the null hypothesis is rejected, many feel that the "best" estimate of u_r is $\bar{x}_r$. We know that the variance of $\bar{x}_r$ is

$$\sigma^2(\bar{x}_r) = \frac{\sigma^2}{n}, \tag{2}$$

if we always use just the sample of n red rulers to estimate u_r. But when the null hypothesis is accepted some people say: "Now that we have accepted the null hypothesis, $u_r = u_b$, the 'best' estimate of u_r is the grand average of the lengths of the $2n$ rulers $(\bar{x}_r + \bar{x}_b)/2$." People with a little different philosophy argue: "No matter how many significance tests you make you can never prove the null hypothesis. You should forget all about the significance test. The 'best' estimate of u_r is $\bar{x}_r$, the mean of the n red rulers." Still others say: "This is a problem of estimation, not one of tests of significance. One should merely inquire what weights to assign to $\bar{x}_r$ and $\bar{x}_b$." No doubt still other people hold still different views.

It may well be that the criterion selected as a significance test and the criterion selected for deciding to pool or not to pool are two different things. For the purpose at hand we shall work with a model for pooling which uses the same test as that given in equation (1). However one may still wish to use a value of k for a significance test which is different from that for the pooling criterion. That is, if one were using a .01 level for a significance test one might require a .20 level, say, before pooling with equal weights.

It should not be thought that the above discussion implies any ambiguity about the meaning of a significance test. Rather, the ambiguity, if it exists, has to do with the use of the usual concept of test of significance as a basis for estimation and action. Of course, the mathematical principles laid down for the use of a significance test and the actual procedure followed often differ. For example, in theory one sets a significance level in advance, one makes a test of significance, and one accepts or rejects the null hypothesis at the significance level previously chosen.

Actual practice usually differs from the procedure outlined above. The person working with the data has no fixed criterion in mind, but rather a vague feeling that the .05 level of significance is pretty widely used, that the .01 level is somewhat more stringent, and that the .001 level is something an experimenter's results do not exceed too often. A standard calculation procedure is now performed. Now the experimenter tends to report the percentage of times his result would have been exceeded had the null hypothesis been true. The words "highly significant" are usually tossed about in connection with results beyond the .001 level, and special remarks are included about results which just do not or barely do exceed the .05 level.

One might argue correctly that those who practice the procedures followed above are not performing significance tests at all, but that they are using the standard procedure to help describe the data and the results of the analysis. Nevertheless, a very large proportion of those who analyze data follow the practice of not setting a significance level in advance (or at most a fuzzy one), and at the same time refer to their procedure as a significance test.

The purpose of the present discussion of theory and practice in connection with significance tests is not meant as a criticism of anyone's procedure. It is presented, because it helps indicate how arguments can arise about the interpretation of significance tests.

PRECISION OF ESTIMATION

First we need a preliminary agreement upon an objective way of measuring which of two estimates is the better. If it is insisted that the estimate of u_r must always be unbiased, proponents of pooling the two samples have already lost their case (so poolers should not be among those who insist on unbiased estimates). However a more temperate view might be that the mean-square deviation D^2 from the true parameter value u_r is a reasonable measure of the goodness of the estimate. We shall use this criterion, although others are equally valid. This criterion has some computational advantages for the models we will construct.

If we always pool the two samples with equal weights, the mean-square deviation of the pooled estimate of u_r, $\bar{x}=(\bar{x}_r+\bar{x}_b)/2$, is

$$D^2(\bar{x}) = \frac{\sigma^2}{2n} + \frac{d^2}{4}, \tag{3}$$

where $d=u_r-u_b$, the difference between the means of the red and black rulers.

On the basis of the mean-square deviation criterion, those who recommend constantly pooling with equal weights have the advantage when

$$\frac{n}{2}$$

while those who recommend using only the measurements from the n red rulers have the advantage when

$$\frac{n}{2} d^2 > \sigma^2.$$

If d^2 is as great as $2\sigma^2$, there is never any advantage in pooling. Furthermore, for any fixed σ and d, d not zero, we can always take samples sufficiently large that there will be no advantage in pooling.

ESTIMATION WITH FIXED DIFFERENCE

We have not done complete justice to those recommending pooling. They point out that they only pool when their criterion suggests that they should. Consequently, we must evaluate the expected value of

$$[\tfrac{1}{2}(\bar{x}_r + \bar{x}_b) - u_r]^2$$

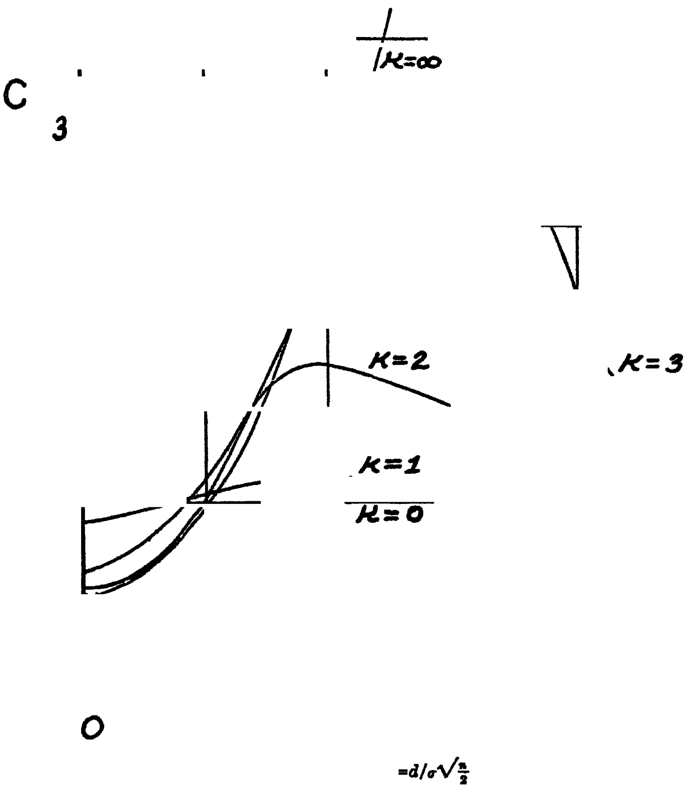
over the acceptance region, plus the expected value of $(\bar{x}_r-u_r)^2$ over the rejection region. When this is done we can divide the mean-square deviation by σ^2/n, the variance when one never pools, and thus get a disadvantage coefficient. Essentially we are taking the variance of the estimate $\bar{x}_r$ as the standard.

The integrations are easy to perform, and the disadvantage coefficient turns out to be

$$C = \frac{1}{2}\left\{1 + x^2 \int_{-k-x}^{k-x} N(y)dy + (k - x)N(k - x) \right.$$
$$\left. + (k + x)N(k + x) + \int_{k-x}^{\infty} N(y)dy + \int_{-\infty}^{-k-x} N(y)dy\right\}.$$

In equation (4), $x = d/\sigma\sqrt{n/2}$, the distance between the true means measured in terms of the standard deviation of the difference of the sample means; k is the critical ratio indicated in equation (1); and $N(U)$ is the normal distribution $e^{-1/2U^2}/\sqrt{2\pi}$.

FIGURE 1



THE DISADVANTAGE COEFFICIENT C AS A FUNCTION OF $x = d\sqrt{n}/\sigma\sqrt{2}$ FOR $k=0, 1, 2, 3, \infty$, WHEN d IS FIXED IF $|t| < k$, $\bar{x} = (\bar{x}_r + \bar{x}_b)/2$ IS USED TO ESTIMATE u_r, WHILE IF $|t| \geq k$ THE ESTIMATE $\bar{x}_r$ IS USED

The behavior of C is more easily understood if one examines the graph shown in Figure 1. We only show positive differences, because the figure is symmetrical about $x=0$. Now providing the difference is very small we can get advantages from pooling up to 10 per cent when $k=1$, up to about 35 per cent when $k=2$, up to 47 per cent when $k=3$, and up to 50 per cent when $k=\infty$. When $k=0$ we never pool—that is, we never expect to find $\bar{x}_r=\bar{x}_b$, exactly. When $k=\infty$, we always pool, that is we set our criterion so large that we invariably pool.

The disadvantage of pooling occurs when x gets large. If we always pool, the mean-square deviation can be so large that the disadvantage coefficient tends to infinity. If we have any criterion which stops us from pooling, it will happen that when the difference between means gets too large we will not pool. Consequently for any finite value of k, as x increases, the disadvantage coefficient will rise to a maximum, and then return asymptotically to $C=1$. It will be noted that when $k=3$ it can happen that the mean-square deviation is about three times the variance we get when we never pool. Similarly for $k=2$, C can grow as large as 1.75, while for $k=1$ the maximum for C is about 1.12.

Diagrams such as Figure 1 suggest that if one has any additional information about d, then something might be gained by pooling or not depending on the indication of the figure. For example, if one were confident that $|d|\leqq 1$, $\sigma=30$, $n=50$, then $x\leqq .167$. In this case one should always pool.

There are examples where the distribution of the true mean can reasonably be thought to be approximately normally distributed. If we examine the day to day performance of an individual on certain psychophysical tests after the initial training period we find that the distribution of his test-to-test means can be approximated by a normal distribution even though his test-to-test variation cannot be explained by random sampling errors from a single population. Also, if we examine the distribution of mean *educational age* of 6th grade classes in a certain area, we will find these means to be approximately normally distributed but with variance too large to be explained by random sampling from the same population.

It is not the intention of the author to argue that there are large numbers of examples of normally distributed means, but rather that there are many examples of cases where the parameter can be thought of as having a distribution of some sort. In such cases there is a possibility of improving an estimate of u_r if it is known that two samples (or more) are drawn one from a population with mean u_r, another from a population with mean u_b, where both u_r and u_b are samples from a parent population of true means.

ESTIMATION WHEN DIFFERENCE IS DISTRIBUTED

We have seen how the disadvantage coefficient C behaves when the difference d between the means is fixed. In a series of experiments, or a lifetime of estimation, we might prefer to think of d as varying. If we knew something about the frequency with which d takes on different values we might be able to take this into account. Of course this is the fundamental Bayes Theorem difficulty. However if we are willing to consider the problem from the point of view of a long series of similar experiments, we can construct some instructive models.

For example, we might be willing to assume that d is normally distributed with zero mean and variance $a^2\sigma^2$, for a series of experiments of a certain type for which we would take samples of the same size n. It is not unreasonable to assume that d is symmetrically distributed about zero, because the choice of the order $u_r - u_b$ or $u_b - u_r$ is arbitrary. The assumption of normality is a more difficult case to argue. It will not be attempted. Instead it is suggested that it would be useful to have a number of additional models of various kinds worked out for problems of this nature.

STANDARD ERROR OF THE ESTIMATE

It needs to be emphasized that in some cases one is anxious to have an accurate estimate and is willing to sacrifice knowledge of its accuracy, while in other cases it is very important to have a "reasonably" unbiased estimate of the standard error of the estimate. If one never pools, one can use σ^2/n as the standard error as usual. In cases where we always pool the standard error of the estimate will be $\sigma^2/2n$, even though the estimate may be biased considerably. When we use a "sometimes pool" procedure the situation is much more complicated. It would, of course, be inaccurate to assign the standard error the value σ^2/n, when we do not pool, and $\sigma^2/2n$ when we do. Instead we must think of the distribution of the estimate of the mean and get its standard error for repeated sampling. The standard error will depend on the significance level and on the true difference between the means. We may underestimate the standard error if we use either of the above formulas.

If it is assumed that d is normally distributed, we may expect the poolers to continue to pool when $|t| < k$, and not to pool when $|t| \geqq k$, making their estimates in the same way as described above.

Again we will compute the disadvantage coefficient C. This time the integrations are more tedious, but we come out with

$$C = 1 + \frac{y-2}{2}\left\{ -k\sqrt{\frac{2}{y+2}}\,N\!\left(k\sqrt{\frac{2}{y+2}}\right) + \int_0^{k\sqrt{2/y+2}} N(v)dv \right\}, \tag{5}$$

where $y = na^2$. It will be remembered that n is the sample size, $a^2\sigma^2$ is the variance of d, and k is the criterion for pooling or not pooling. As before $N(U)$ is the normal distribution with zero mean and unit variance.

An examination of equation (5) shows that when $k=0$, $C=1$, which is what we expect if we never pool. Further, when $k=\infty$ we always pool, and the limit $C=(2+na^2)/4$, which can easily be checked to be the expected value of $D^2(\bar{x})/\sigma^2(\bar{x}_r)$ when d has the distribution we have assigned it.

It can be shown that when y is large C tends to unity asymptotically from above like $1+k^3\sqrt{2}/3\sqrt{2\pi y}(1+5/y)$. This means once again that if a pooler almost always gets large discrepancies in the t-test, he will only pool occasionally, and the effect will be essentially as if he did not pool at all.

Figure 2 shows more clearly how C behaves as $\sqrt{y}=a\sqrt{n}$ goes from 0 to 6. We see that all the curves for $0<k<\infty$ are contained between the curve $C=(2+y)/4$ for $k=\infty$ and the straight line $C=1$ for $k=0$. It can be verified from equation (5) that when $\sqrt{y}=\sqrt{2}$, $C=1$.

With the present model we note that for small values of $na^2=y$ there are advantages up to 50 per cent to be gained by pooling. For $na^2=y>2$ some loss in precision will occur. However, with the present model the losses do not go as high as when we dealt with a fixed value of d (Fig. 1), providing $k<\infty$.

For convenience, we have graphed C against $\sqrt{y}$, for $0 \leqq \sqrt{y} \leqq 6$. For $k=3$ a very flat maximum occurs at about $\sqrt{y}=4.5$, and $C=1.7$; for $k=2$ the maximum occurs at about $\sqrt{y}=4$ and $C=1.2$; while for $k=1$, the maximum occurs at about $\sqrt{y}=3$, and $C=1.04$.

We compare the maximum C's for the two models discussed in Table I.

Of course, even if one accepts the present model, one still has to make a decision between pooling and not pooling. We do not usually enter into an important experiment with no ideas at all about the order of magnitude of the quantities we will find.

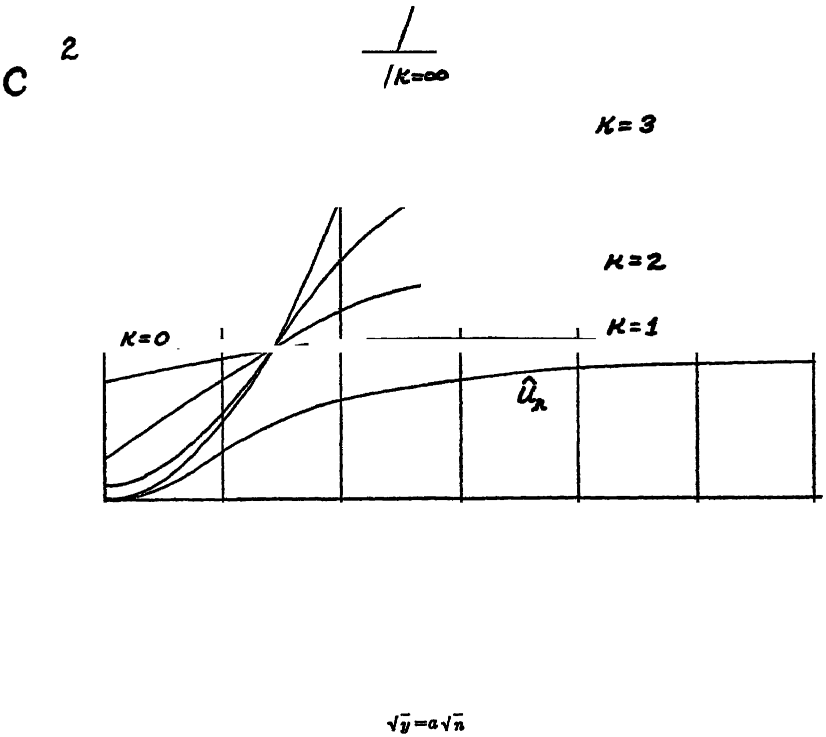
TABLE I

MAXIMUM DISADVANTAGE COEFFICIENT C IN POOLING FOR TWO CASES: 1) FIXED DIFFERENCE BETWEEN MEANS d; 2) d IS NORMALLY DISTRIBUTED. THE CRITERION FOR POOLING OR NOT POOLING IS k

k	d *fixed*	d *normal*
0	1.00	1.00
1	1.12	1.04
2	1.75	1 2
3	3.0	1 7

For example, suppose we believe that the standard deviation of the difference between the mean lengths of successive pairs of shipments of red and black rulers is about .001 inch, that the standard deviation of red and black rulers is about .010 inch. If we take a sample of 50

FIGURE 2



THE DISADVANTAGE COEFFICIENT C AS A FUNCTION OF $\sqrt{y}=a\sqrt{n}$ FOR $k=0$, 1, 2, 3, ∞, WHEN d IS NORMALLY DISTRIBUTED WITH ZERO MEAN AND UNIT VARIANCE. IF $|t|<k$, $\bar{x}=(\bar{x}_r+\bar{x}_b)/2$ IS USED TO ESTIMATE u_r, WHILE IF $|t|\geqq k$, THE ESTIMATE $\bar{x}_r$ IS USED. THE LOWEST CURVE, $\hat{u}_r$, IS BASED ON THE MAXIMUM LIKELIHOOD ESTIMATE $\hat{u}_r=(\bar{x}_r(1+y)+\bar{x}_b)/(2+y)$.

rulers of each color, then $a = .001/.010 = 0.1$, and $a\sqrt{n} = .707$. So no matter whether we are using $k = 1$, 2, or 3 as our criterion, Figure 2 indicates that the advantage is "probably" with the poolers. The word "probably" is used because the quantity .001 inch given above is really a guess on the part of the experimenter. There is nothing reprehensible about intelligent guessing. The most successful experimenters are those who combine a good knowledge of their subject matter and of experimental design with ability to make judicious guesses. The efficiency of experiments can almost always be improved by good guessing.

On the basis of this model, therefore, if there is some possibility of making intelligent guesses about a and σ, the author would decide between pooling and not pooling on the basis of C and the sample. In the absence of such guesses the author feels that between those who never pool, and those who require a two-sigma agreement ($|t| < 2$) or better before pooling, there is little to choose. It does not seem likely that the latter can be injured badly on the average, and they may gain considerably.

On the basis of the present investigation the author's principal reason for disliking the use of a k as liberal as 3 comes from the disadvantage factors when d is regarded as fixed. It is always possible for an experimenter to encounter a series of experiments where d does not behave as it is assumed to do in the variable model, and in that case one may very well be multiplying the no-pooling precision by a factor of 3.

Of course, if the observations are cheap and easy to get, the problem scarcely arises; one merely takes a larger sample of red rulers.

MAXIMUM LIKELIHOOD POOLING

We have gone to a good deal of trouble to inquire about what happens when the decision to pool or not to pool is either made a priori, or made on the basis of a criterion which is equivalent to a significance test. We have kept to this line because we were interested in this question, and because this is a method often used in pooling.

An important model which might be mentioned however is the maximum likelihood method. If we consider the case of d fixed, the maximum likelihood estimate is $\bar{x}_r$, and $C = 1$, as we have noted. But if we assume as in the previous model that d is normally distributed, the maximum likelihood estimate suggests that we always pool, but assign varying weights to $\bar{x}_r$ and $\bar{x}_b$. The maximum likelihood estimate of u_r is just

$$(6) \qquad \hat{u}_r = \frac{\bar{x}_r(1 + na^2) + \bar{x}_b}{2 + na^2}.$$

If $na^2 = 0$, that is $d = u_r - u_b = 0$, the pooled estimate $(\bar{x}_r + \bar{x}_b)/2 = \bar{x}$ is the maximum likelihood estimate. If na^2 is extremely large compared to $\bar{x}_r$, $\bar{x}_b$, and 2, then the maximum likelihood estimate reduces to $\bar{x}_r$.

The mean-square deviation of $\hat{u}_r$ from u_r, using the model with d normally distributed gives

$$D^2(\hat{u}_r) = \frac{\sigma^2(1 + na^2)}{n(2 + na^2)} \tag{7}$$

Here we see that as na^2 goes from zero to infinity the pooled estimate has mean-square deviation running from $\sigma^2/2n$ to σ^2/n, that is from the variance for $2n$ observations to the variance for n observations. The lowest curve (labelled $\hat{u}_r$) of Figure 2 is a graph of $D^2(\hat{u}_r)$.

Equation (7) is also the asymptotic variance of $\hat{u}_r$ given by maximum likelihood methods when the true value of a^2 is unknown, but to be estimated from $\bar{x}_r$ and $\bar{x}_b$. It is not clear however exactly what estimate of a^2 should be used for small samples. However, the estimate is approximately $(\bar{x}_r - \bar{x}_b)^2/\sigma^2$. If this value is used for a^2 in equation (6) it is clear that when the difference between the sample means is large $\bar{x}_r$ gets most of the weight, while if it is small the weights are more equally balanced between the sample means.

The curve for $\hat{u}_r$ in Figure 2 shows that the disadvantage factor for the maximum likelihood method is always less than unity. This implies that if one has reason to believe that d is normally distributed the maximum likelihood solution is by far the most profitable among the ones discussed. The implication is that at least for this situation pooling with unequal weights is usually the best policy. The all-or-none criterion method cannot compare with the estimate of equation (6).

This result is rather important because it backs up our intuitive feeling that we should use all the information at hand in making our estimates.

CONCLUSION

We have presented and discussed several ways of pooling data from two samples to estimate the mean of the population of one of them. Perhaps the most important point of the discussion is that for many kinds of estimation, not only of means, we need more models to study. These models help us understand what is going on when we pool data from several samples to make a single estimate. Furthermore they may assist us in deciding which of several possible methods of combining data may be preferable.

The author is indebted to J. Wolfowitz for calling his attention to an error in an early draft of this paper.

BIBLIOGRAPHY

T. A. Bancroft, "On Biases in Estimation Due to the Use of Preliminary Tests of Significance," *Annals of Mathematical Statistics*, Vol. 15, 1944, pp. 190–204.

Joseph Berkson, "Tests of Significance Considered as Evidence," *Journal of the American Statistical Association*, Vol. 37, 1942, pp. 324–335.

USE OF VARIANCE COMPONENTS IN THE ANALYSES OF HOG PRICES IN TWO MARKETS

R. L. ANDERSON
Institute of Statistics, North Carolina State College

Corrections to article published in Volume 42, No. 240 December 1947

Page 618, Table 4: The variance components for the *days* mean square should include $24\sigma^2_{dy}$.

Page 624, line 2 should read: $\doteq V_{dmy}+V_{myc}-V_{dmyc}$

Page 624, line 4 above the table should read:

$$M: \quad V_{mc}+V_{dmy}-V_{dmyc}.$$

Page 625 in the Table for F', last line, should read: 3.34*.

Page 360, line 6 from the bottom should read: $a_j=\pm 1$.

ON ESTIMATING PRECISION OF MEASURING INSTRUMENTS AND PRODUCT VARIABILITY

FRANK E. GRUBBS
Ballistic Research Laboratories, Aberdeen Proving Ground, Md.

A measurement or observed value is discussed as the sum of two components—one the absolute value of the characteristic measured and the other an error of measurement. The variation in absolute values of the characteristic or items measured is termed product variability, whereas the variation in errors of measurement of an instrument is called the precision or reproducibility of measurement. Techniques are given for separating and estimating product variability and precision of measurement. Comparisons of the various techniques are also discussed for cases involving two or more instruments.

1. INTRODUCTION

A MEASUREMENT consists of the sum of two components, one component being the true or absolute value of the item or characteristic measured and the other an error of measurement. In making measurements, the aim is to obtain as nearly as practicable the true or absolute value of every characteristic measured. This is accomplished either by attempting substantially to eliminate errors of measurement or by using an instrument for which the errors of measurement are known to a sufficient or estimable degree. In particular, if the errors of measurement are for all practical purposes constant then rather accurate knowledge of the true values of the items measured can be ascertained when the size of the constant error can be calibrated by reference to a standard. On the other hand, if the errors of measurement vary appreciably from one observation to another in a series of observations, then the accuracy and usefulness of the reported results will depend largely on available knowledge concerning the relation between the errors of measurement and the true values of the characteristics or items measured. Thus, as the general demand for closer tolerances increases it becomes important to investigate methods for studying the relation between the two components of measurement in order to insure that the instrumentation involved meets its purpose adequately.

The purpose of the present investigation, therefore, is to present a statistical study of methods for estimating and comparing product variations and errors of measurement. The methods discussed are quite general and should apply to most instrumentation set-ups which can be employed as outlined below. The problem itself arose at Aberdeen Proving Ground in connection with investigations of the adequacy of electric clocks in measuring burning times of powder train fuzes, in measurements of velocity with two or more chronographs, and in measures of rotating band clearance, to mention some.[1]

2. MEASUREMENTS WITH A SINGLE INSTRUMENT

Suppose a single instrument is used to measure some characteristic of n different items. The results of measurement may be represented symbolically as follows:

$$\begin{array}{c} x_1 + e_1 \\ x_2 + e_2 \\ \cdot \\ \cdot \\ \cdot \\ x_i + e_i \\ \\ \\ x_n + e_n \end{array}$$

where $x_1, x_2, \cdots x_1, \cdots x_n$ are the true values of the characteristics or the items measured and the quantities, $e_1, e_2, \cdots e_i, \cdots e_n$ are errors of measurement. Generally speaking, $x_1, x_2, \cdots x_i, \cdots x_n$ are all different since there is variability in the items or product measured. Also, the quantities e_1, e_2, etc. will usually vary because a different error in measurement may be made on each item. It is to be noted here that in the single measurement or observed value, $x_i + e_i$, the error of measurement, e_i, is inseparable from the true value of the characteristic, x_i. That is, for example, a single micrometer measurement may be, say, .5635 inch and it is not known precisely what part of .5635 inch is the error of measurement, which in this case arises from the combination consisting of the operator and the micrometer. If all the errors of measurement e_1, e_2, etc. were zero it would be said that the instrument is accurate or perfectly calibrated for each measurement (or that the measurements are absolute) and also that the utmost in precision or reproducibility has been attained. Precision here is defined as a measure of the variation in the errors of measurement,

---

[1] Somewhat similar problems, namely those of estimating components of variance, are treated in Snedecor's book *Statistical Methods*, Iowa State College Press, under the index heading "Variance, Components of" and in a paper by H. E. Daniels "The Estimation of Components of Variance," *Journal of the Royal Statistical Society*, Supp., Vol. VI, No. 2, 1939, pp. 186–197.

e_1, e_2, etc. If the quantities e_1, e_2, etc. vary considerably it may be said that the measuring instrument lacks precision. On the other hand, if all of the errors of measurement are equal, i.e. if $e_1 = e_2 = \cdots = e_i = \cdots e_n$, our instrument possesses the utmost in *precision* because the same constant error in measurement is made each time. However, such an instrument does not give absolute measurements unless corrections are made for the constant error of measurement. It is seen, therefore, that accurate knowledge of the true values or characteristics, x_1, x_2, x_3, etc. cannot be obtained unless something is known about the errors of measurement, e_1, e_2, etc.

If the measured or observed values $x_1 + e_1$, $x_2 + e_2$, etc. are averaged, one obtains

$$\frac{\sum_{i=1}^{n} (x_i + e_i)}{n} = \bar{x} + \bar{e}$$

where $\bar{x}$ is the average value of the characteristic measured and $\bar{e}$ is the average error of measurement. (As in the case of $x_i + e_i$ above, the average observed value, $\bar{x} + \bar{e}$, consists of inseparable components also.) If actually, $\bar{e} = 0$, then it may be said that the errors of measurement average out to zero and that the average observed value would be "accurate"; but on the other hand, the measurements may not be *precise* since the individual errors of measurement may vary considerably from one observation to another.

In order to determine the precision or reproducibility of measurement, the amount of variation in the quantities e_1, e_2, etc. should be examined, especially as compared to the variation in $x_1, x_2, \cdots x_n$. Now, the variance of the observed measurements is given by

$$\begin{aligned}
&\frac{1}{n-1} \sum_{i=1}^{n} \left\{ (x_i + e_i) - (\bar{x} + \bar{e}) \right\}^2 \\
&\quad = \frac{1}{n-1} \sum_{i=1}^{n} \left\{ (x_i - \bar{x}) + (e_i - \bar{e}) \right\}^2 \\
&\quad = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 + \frac{2}{n-1} \sum_{i=1}^{n} (x_i + \bar{x})(e_i - \bar{e}) \\
&\qquad + \frac{1}{n-1} \sum_{i=1}^{n} (e_1 - \bar{e})^2.
\end{aligned}$$

The quantity

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

which is based on a sample of n measurements, gives an unbiased estimate ($n-1$ degrees of freedom used) of the true variation or variance in the characteristic x of a larger category from which the random sample of n observations is taken. In a like manner,

$$s_e^2 = \frac{1}{n-1}\sum_{i=1}(e_i - \bar{e})^2$$

is a measure of the variation in the errors of measurement and gives an unbiased estimate of the variance in errors of the measuring instrument. The quantity,

$$s_{xe} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(e_i - \bar{e}),$$

known as a co-variance, gives a measure of the relation between the values of the characteristics measured and the errors of measurement. In this discussion, it will be assumed that over a somewhat limited range in values of the characteristic, x, no correlation exists between the values of the characteristic measured and the errors of measurement, and consequently that the mean or expected value of the co-variance, s_{xe}, will be zero (see below). It may be true that errors of measurement in some cases increase with increasing magnitude of the characteristic measured; however, in many problems the x and e are sufficiently independent to insure that limited variations in x are not reflected in the errors of measurement.

If s_e^2 is small (on the average), relatively speaking, then it may be said that the measuring instrument is precise; on the other hand, the larger s_e^2, the more imprecise are the measurements. If s_x^2 is small, there is little variation in the characteristic measured: otherwise the variability in the characteristic measured or product is not small. The efficiency with which the absolute values $x_1, x_2, \cdots, x_n$ are measured depends on (1) the relative size of s_e^2 and s_x^2, the two components it is desired to estimate and (2) proper calibration of the measuring instrument. An important point of interest here is that in using a single measuring instrument on a group of items, the variance of the observations consists of the sum of *inseparable* components—one the variance in the items measured and the other the variance in the errors of measurement.

With a single measuring instrument, therefore, it apparently is not possible (1) to separate an individual error of measurement, e_i, from the value of the characteristic, x_i, (2) to separate the *average* error of measurement, $\bar{e}$, from $\bar{x}$, the average value of the characteristic measured and (3) to separate the variance in errors of measurement from the variance in the characteristics measured in order that the relative magnitude of these components could be studied. It is true that by measuring the same characteristic over and over again without bias or if there is no variation in the product (i.e. if $x_1 = x_2 = \cdots = x_n$), then s_e^2 could be obtained directly. Also, $\bar{e}$ could be determined if the value of x were known or if the instrument itself were calibrated accurately against a standard.

The usefulness of a measuring instrument depends therefore on knowledge concerning the errors of measurement and a factor of primary importance is the relative order of magnitude of s_x^2 and s_e^2. Thus, methods are sought for separating the components, s_x^2 and s_e^2. The actual separation of s_x^2 and s_e^2 requires the use of at least two instruments as pointed out below. First, however, certain definitions and assumptions will be given.

3. DEFINITIONS

Let x_i = the absolute measurement or true value of the characteristic considered, where i represents the ith item in the sample

e_{ij} = the error in measurement of the ith item by the jth instrument, I_j

The observed values are then $x_i + e_{ij}$ and the mean measurement,

$$\bar{x} + \bar{e}_j = \frac{1}{n} \sum_{i=1}^{n} (x_i + e_{ij})$$

Taking

$$s_{uv} \quad \text{or} \quad s_{u,v} = \frac{1}{n-1} \sum_{i=1}^{n} (u_i - \bar{u})(v_i - \bar{v}),$$

it follows that

$$s_u^2 = s_{uu} = \frac{1}{n-1} \sum_{i=1}^{n} (u_i - \bar{u})^2.$$

Define also the following expected values:

$$E(x_i) = \mu, \qquad E(e_{ij}) = \epsilon_j$$

$$E(x_i - \mu)^2 = \sigma_x^2, \qquad E(e_{ij} - \epsilon_j)^2 = \sigma_{e_j}^2.$$

Hence,

$$E(s_{uu}) = \sigma_u^2.$$

4. ASSUMPTIONS

We will assume that x_i and e_{ij} are both normally distributed (although this is not necessary in so far as estimating variance only is concerned) and that the absolute values of the characteristic measured and the errors of measurement are statistically independent, i.e.

$$E(x_i^p \cdot e_{ij}^q) = E(x_i^p) \cdot E(e_{ij}^q).$$

In addition, it will be assumed that the errors of measurement of the different instruments are also statistically independent of each other, i.e.

$$E(e_{ij}^p \cdot e_{ik}^q) = E(e_{ij}^p) \cdot E(e_{ik}^q) \qquad j \neq k.$$

5. THE SEPARATION OF s_x^2 AND $s_{e_j}^2$ WITH TWO INSTRUMENTS

If two instruments I_1 and I_2 are used to make measurements on the same n items, the observed values can be represented symbolically as follows:

Measurements by I_1	Measurements by I_2
$x_1 + e_{11}$	$x_1 + e_{12}$
$x_2 + e_{21}$	$x_2 + e_{22}$
$x_i + e_{i1}$	$x_i + e_{i2}$
$x_n + e_{n1}$	$x_n + e_{n2}$

For the case of two instruments, there are two methods or computational procedures for estimating precision of measurement.

Method A_2:

As an estimate of the variance in errors of measurement of instrument I_1, we take[2]

$$\text{est. } (\sigma^2_{e_1}) = s^2{}_{x+e_1} - s_{x+e_1,x+e_2} \tag{1}$$

where

$$\begin{aligned} s_{x+e_1,x+e_2} &= \frac{1}{n-1}\sum_{i=1}^{n} \{(x_i + e_{i1}) - (\bar{x} + \bar{e}_1)\}\{(x_i + e_{i2}) - (\bar{x} + \bar{e}_2)\} \\ &= \frac{1}{n(n-1)}\left\{n\sum_{i=1}^{n} (x_1 + e_{i1})(x_{i1} + e_{i2}) \right. \\ &\qquad \left. - \left[\sum_{i=1}^{n} (x_i + e_{i1})\right]\left[\sum_{i=1}^{n} (x_i + e_{i2})\right]\right\} \end{aligned}$$

and as an estimate of the variance in errors of measurement of instrument I_2, we compute

$$\text{est. } (\sigma_{e_2}{}^2) = s^2{}_{x+e_2} - s_{x+e_1,x+e_2}. \tag{2}$$

The estimates (1) and (2) are unbiased, since, for example,

$$\text{est. } (\sigma_{e_1}{}^2) = s_{e_1}{}^2 + s_{xe_1} - s_{xe_2} - s_{e_1e_2} \text{ symbolically}$$

and

$$E(s_{e_1}{}^2) = \sigma_{e_1}{}^2,$$

whereas

$$E(s_{xe_1}) = E(s_{xe_2}) = E(s_{e_1e_2}) = 0 \text{ under the assumptions.}$$

Method B_2:

It is remarked in passing that a technique of working with differences in corresponding measurements of the two instruments is of some interest. That is, if the measurements of I_2 are substracted from corresponding measurements of I_1, the result gives

$$\begin{array}{c} e_{11} - e_{12} \\ e_{21} - e_{22} \\ \\ \\ e_{i1} - e_{i2} \\ \\ \\ e_{n1} - e_{n2}, \end{array}$$

---

[2] Estimates (1) and (2) are maximum likelihood estimates.

thus eliminating the true value of x_i, the characteristic measured. Using the above differences, an estimate of $\sigma_{e_1}^2$ is given by

(3) $$\text{est. } (\sigma_{e_1}^2) = \tfrac{1}{2}\left\{s^2_{x+e_1} - s^2_{x+e_2} + s^2_{e_1-e_2}\right\}.$$

It is easy to see that estimate (3) is symbolically

$$s_{e_1}^2 + s_{xe_1} - s_{xe_2} .$$

and hence precisely that given by (1).

In order to estimate the amount of variability in the product being measured, we may first add corresponding measurements of the two instruments, obtaining symbolically

$$\begin{array}{c} x_1 + e_{11} + x_1 + e_{12} \\ \vdots \\ x_i + e_{i1} + x_i + e_{i2} \\ \vdots \\ x_n + e_{n1} + x_n + e_{n2}. \end{array}$$

The variance of the above sums of readings is symbolically

$$4s_x^2 + s_{e_1}^2 + s_{e_2}^2 + 4s_{xe_1} + 4s_{xe_2} + 2s_{e_1e_2}$$

so that an estimate of the variability in the product, σ_x^2, is given by

(4) $$\text{est. } (\sigma_x^2) = \tfrac{1}{4}\left\{s^2_{x+e_1+x+e_2} - s^2_{e_1-e_2}\right\}.$$

Actually, however, the estimate of σ_x^2 is given simply by the covariance of the readings of I_1 and I_2:

(4a) $$\text{est. } (\sigma_x^2) = s_{x+e_1,x+e_2} \text{ (4 and 4a are identical).}$$

Estimate (4) or (4a) is obviously unbiased and can be shown to be the maximum likelihood estimate.

In order to determine the reliability of our estimates of precision in measurement, we examine the variances of the estimates. Now the population variance of (1) is given by

$$E\left\{\text{est. } (\sigma_{e_1}^2) - \sigma_{e_1}^2\right\}^2 = E\left\{s_{e_1}^2 + s_{xe_1} - s_{xe_2} - s_{e_1e_2}\right\}^2 - \sigma_{e_1}^4.$$

It is well-known[3] that we may write

$$E(s_{e_1}{}^4) - \sigma_{e_1}{}^4 = \frac{1}{n(n-1)}\left\{(n-1)\mu_{4:e_1} + (3-n)\sigma_{e_1}{}^4\right\}$$
$$= \frac{2}{n-1}\,\sigma_{e_1}{}^4$$

for a normal parent where $\mu_{4\cdot e_1}$ is the fourth central moment of e_{i1}. Also

$$E(s_{xe_1}{}^2) = \quad 1$$

$$E(s_{xe_2}{}^2) = \frac{1}{n-1}\,\sigma_x{}^2\sigma_{e_2}{}^2,$$

$$E(s_{e_1e_2}{}^2) = \frac{1}{n-1}\,\sigma_{e_1}{}^2\sigma_{e_2}{}^2$$

so that

$$E\left\{\text{est. }(\sigma_{e_1}{}^2) - \sigma_{e_1}{}^2\right\}^2$$
$$= \frac{2}{n-1}\,\sigma_{e_1}{}^4 + \frac{1}{n-1}\,(\sigma_x{}^2\sigma_{e_1}{}^2 + \sigma_x{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_2}{}^2). \tag{5}$$

Likewise, the population variance of the estimate of $\sigma_{e_2}{}^2$ is given by

$$E\left\{\text{est. }(\sigma_{e_2}{}^2) - \sigma_{e_2}{}^2\right\}^2$$
$$= \frac{2}{n-1}\,\sigma_{e_2}{}^4 + \frac{1}{n-1}\,(\sigma_x{}^2\sigma_{e_1}{}^2 + \sigma_x{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_2}{}^2) \tag{6}$$

and the population variance of the estimate of product variability is given by

$$E\left\{\text{est. }(\sigma_x{}^2) - \sigma_x{}^2\right\}^2$$
$$= \frac{2}{n-1}\,\sigma_x{}^2 + \frac{1}{n-1}\left\{\sigma_x{}^2\sigma_{e_1}{}^2 + \sigma_x{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_2}{}^2\right\}. \tag{7}$$

It is to be noted that the variance of the estimate, est. $(\sigma_{e_1}{}^2)$ or est. $(\sigma_{e_2}{}^2)$ depends on (a) $\sigma_x{}^2$, the variance in the characteristic measured,

---

[3] See, for example, R. A. Fisher, "Moments and Product Moments of Sampling Distributions." *Proceedings of the London Mathematical Society*, Vol. 30, 1929, pp 215–218.

(*b*) $\sigma_{e_1}^2$, the variance of the errors of measurement of instrument I_1, (*c*) $\sigma_{e_2}^2$, the variance of the errors of measurement of instrument I_2 and (*d*) n, the number of observations or the sample size. It is seen, therefore, that in order to obtain a precise estimate of $\sigma_{e_1}^2$ when using only two instruments, the variation in the characteristic measured, i.e. σ_x^2, should be held to a reasonable minimum or the sample size, n, should be sufficiently large.

If the variation in the characteristic measured is zero (or if we measure the same item over and over again), i.e. if $\sigma_x^2=0$, then one could compute

$$\frac{1}{n-1}\sum_{i=1}^{n}(e_{i1}-\bar{e}_1)^2$$

directly, and this would give an estimate of $\sigma_{e_1}^2$ with variance equal to

$$\frac{2}{n-1}\sigma_{e_1}^4.$$

Apparently, in employing two instruments, there are only two computational procedures of interest for separating the variability in the product from the variance in the errors of measurement and both methods give the same estimate. In using either method, however, it is possible to estimate $\sigma_{e_1}^2$, $\sigma_{e_2}^2$ and σ_x^2 and thus determine from the relative order of magnitude of these quantities whether the instruments are sufficiently precise to use in taking the required measurements.

It is pointed out that the average observed value of I_1 is $\bar{x}+\bar{e}_1$ and the average reading of I_2 is $\bar{x}+\bar{e}_2$, so that upon subtracting, $\bar{e}_1-\bar{e}_2$ is obtained, the average difference in the errors of measurement of I_1 and I_2. Of course, $\bar{e}_1$ and $\bar{e}_2$ are inseparable here also and estimates of either depend upon calibration of I_1, I_2 or both against a standard.

6. THE SEPARATION OF s_x^2 AND $s_{e_j}^2$ WITH THREE INSTRUMENTS[4]

If three instruments I_1, I_2 and I_3 are used to make measurements on the same n items, the observed values can be represented as:

---

[4] The questions of optimum estimates and efficiency of methods suggested below for three or more instruments are not answered in this paper. It is believed, however, that the suggested techniques show promise of providing good, workable estimates. It is hoped that a more complete investigation of this problem will be undertaken in the near future and that others will interest themselves in this rather important study.

Measurements by I_1	Measurements by I_2	Measurements by I_3
$x_1 + e_{11}$	$x_1 + e_{12}$	$x_1 + e_{13}$
$x_2 + e_{21}$	$x_2 + e_{22}$	$x_2 + e_{23}$
$\vdots$	$\vdots$	$\vdots$
$x_n + e_{n1}$	$x_n + e_{n2}$	$x_n + e_{n3}$

For the case of three instruments two techniques will be considered for estimating $\sigma_{e_1}^2$, $\sigma_{e_2}^2$, and $\sigma_{e_3}^2$, the variances in the errors of measurement of instruments I_1, I_2, and I_3.

Method A_3:

Two computational procedures giving rise to the same estimate will be given here.

First, an estimate of $\sigma_{e_1}^2$ is given by

$$\text{est. } (\sigma_{e_1}^2) = s^2_{x+e_1} - s_{x+e_1,x+e_2} - s_{x+e_1,x+e_3} + s_{x+e_2,x+e_3}. \tag{8}$$

Also, in some particular applications it may be enlightening to eliminate the value of the characteristic measured by working with differences. That is, for each item measured, form the three columns of differences in observed values of the instruments taken two at a time. Thus,

Differences in Readings, $I_1 - I_2$	Differences in Readings, $I_1 - I_3$	Differences in Readings, $I_2 - I_3$
$e_{11} + e_{12}$	$e_{11} - e_{13}$	$e_{12} - e_{13}$
$e_{21} - e_{22}$	$e_{21} - e_{23}$	$e_{22} - e_{23}$
$e_{n1} - e_{n2}$	$e_{n1} - e_{n3}$	$e_{n2} - e_{n3}$

The variances of these three columns will be symbolically

$$s^2_{e_1-e_2} = s_{e_1}^2 - 2s_{e_1e_2} + s_{e_2}^2, \qquad s^2_{e_1-e_3} = s_{e_1}^2 - 2s_{e_1e_3} + s_{e_3}^2,$$
$$s^2_{e_2-e_3} = s_{e_2}^2 - 2s_{e_2e_3} + s_{e_3}^2.$$

Hence, as an estimate of $\sigma_{e_1}^2$, take

$$\text{est. } (\sigma_{e_1}^2) = \tfrac{1}{2}\left\{s^2_{e_1-e_2} + s^2_{e_1-e_3} - s^2_{e_2-e_3}\right\}. \tag{9}$$

Estimates (8) and (9) are identical, both being equal to

$$s_{e_1}^2 - s_{e_1e_2} - s_{e_1e_3} + s_{e_2e_3} \text{ symbolically.}$$

Estimates (8) and (9) are unbiased and also free of the characteristic measured.

As an estimate of $\sigma_{e_2}^2$, form

$$\text{(10)} \qquad \text{est. } (\sigma_{e_2}^2) = s^2_{x+e_2} - s_{x+e_1,x+e_2} + s_{x+e_1,x+e_3} - s_{x+e_2,x+e_3}$$

or carry out a computation similar to (9).

As an estimate of $\sigma_{\epsilon_3}^2$, form

$$\text{(11)} \qquad \text{est. } (\sigma_{e_3}^2) = s^2_{x+e_3} + s_{x+e_1,x+e_2} - s_{x+e_1,x+e_3} - s_{x+e_2,x+e_3}$$

or carry out a computation similar to (9).

The variance of estimate (8) or (9) is given by

$$\begin{aligned} E\{\text{est. } (\sigma_{e_1}^2) - \sigma_{e_1}^2\}^2 &= E\{\text{est. } (\sigma_{e_1}^2)\}^2 - \sigma_{e_1}^4 \\ &= E\{s_{e_1}^2 - s_{e_1e_2} - s_{e_1e_3} + s_{e_2e_3}\}^2 - \sigma_e^4. \end{aligned}$$

In an analogous manner to that for determining the variance of an estimate in the case of two instruments, it is easily seen that for three instruments

$$\text{(12)} \qquad \begin{aligned} &E\{\text{est. } (\sigma_{e_1}^2) - \sigma_e^2\}^2 \\ &\quad = \frac{2}{n-1}\sigma_{e_1}^4 + \frac{1}{n-1}(\sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_3}^2). \end{aligned}$$

Likewise, the variance of the estimate (10) is

$$\text{(13)} \qquad \frac{2}{n-1}\sigma_{e_2}^4 + \frac{1}{n-1}(\sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_3}^2)$$

and the variance of (11) is given by

$$\text{(14)} \qquad \frac{2}{n-1}\sigma_{e_3}^4 + \frac{1}{n-1}(\sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_3}^2).$$

Thus, the variances of the estimates of variation in errors of measurement obtained above for three instruments are entirely free of σ_x^2, the variability in the characteristic measured. Apparently, at least three instruments are needed to eliminate the characteristic x.

Estimate (8) using three instruments will be more precise than estimate (1) employing two instruments provided

$$\sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_3}^2 < \sigma_x^2\sigma_{e_1}^2 + \sigma_x^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_2}^2$$

i.e. if $\sigma_{e_3}^2 < \sigma_x^2.$

Since in general the variance in errors of measurement should be smaller than the variance in absolute values of the characteristic measured, it would appear that there is some advantage in the use of the additional measuring instrument, I_3 for the purpose of estimating $\sigma_{e_1}^2$, $\sigma_{e_2}^2$, or $\sigma_{e_3}^2$.

Method B_3:[5]

This technique involves the variance of the readings of I_1 and the average of the three co-variances which can be formed from the three sets of observed values for instruments I_1, I_2, and I_3.

Now the co-variance of the I_1 and I_2 readings is

$$s_{x+e_1,x+e_2} = s_x^2 + s_{xe_1} + s_{xe_2} + s_{e_1e_2} \text{ symbolically.}$$

The co-variance of the I_2 and I_3 readings is symbolically

$$s_x^2 + s_{xe_2} + s_{xe_3} + s_{e_2e_3}$$

and that of the I_1 and I_3 readings is

$$s_x^2 + s_{xe_1} + s_{xe_3} + s_{e_1e_3}.$$

The average of these three co-variances is symbolically

$$s_x^2 + \tfrac{2}{3}s_{xe_1} + \tfrac{2}{3}s_{xe_2} + \tfrac{2}{3}s_{xe_3} + \tfrac{1}{3}s_{e_1e_2} + \tfrac{1}{3}s_{e_2e_3} + \tfrac{1}{3}s_{e_1e_3}.$$

As previously stated, the variance of the observations with instrument I_1 is symbolically

$$s^2_{x+e_1} = s_x^2 + 2s_{xe_1} + s_{e_1}^2.$$

The latter variance minus the average co-variance gives as an estimate of $\sigma_{e_1}^2$

$$\begin{aligned} \text{est. } (\sigma_{e_1}^2) &= s^2_{x+e_1} - \tfrac{1}{3}\{s_{x+e_1,x+e_2} + s_{x+e_1,x+e_3} + s_{x+e_2,x+e_3}\} \\ &= s_{e_1}^2 + \tfrac{4}{3}s_{xe_1} - \tfrac{2}{3}s_{xe_2} - \tfrac{2}{3}s_{xe_3} - \tfrac{1}{3}s_{e_1e_2} - \tfrac{1}{3}s_{e_1e_3} \\ &\quad - \tfrac{1}{3}s_{e_2e_3} \text{ symbolically} \end{aligned} \tag{15}$$

and the variance of this estimate is, in a manner analogous to that previously indicated, given by

$$\frac{2}{n-1}\sigma_{e_1}^4 + \frac{1}{n-1}\{\tfrac{16}{9}\sigma_x^2\sigma_{e_1}^2 + \tfrac{4}{9}\sigma_x^2\sigma_{e_2}^2 + \tfrac{4}{9}\sigma_x^2\sigma_{e_3}^2 + \tfrac{1}{9}\sigma_{e_1}^2\sigma_{e_2}^2 + \tfrac{1}{9}\sigma_{e_1}^2\sigma_{e_3}^2 + \tfrac{1}{9}\sigma_{e_2}^2\sigma_{e_3}^2\}. \tag{16}$$

---

[5] Mr. H A. Noble of the Ballistic Research Laboratory, using correlation coefficients, proposed a technique for estimating $\sigma_{e_1}^2$ precisely equivalent to Method B_3 and also suggested the problem of comparing the variance of Method B_3 with that of Method A_3 when three instruments are employed.

Recalling that Method A_3 for the case of three instruments gave an estimate whose variance is given by

$$\frac{2}{n-1}\sigma_{e_1}{}^4 + \frac{1}{n-1}(\sigma_{e_1}{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_3}{}^2 + \sigma_{e_2}{}^2\sigma_{e_3}{}^2),$$

it may be seen that Method B_3 will not give an estimate of $\sigma_{e_1}{}^2$ as precise as that of Method A_3 unless

$$\sigma_x{}^2 < \frac{2(\sigma_{e_1}{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_3}{}^2 + \sigma_{e_2}{}^2\sigma_{e_3}{}^2)}{4\sigma_{e_1}{}^4 + \sigma_{e_2}{}^2 + \sigma_{e_3}{}^2}\,.$$

Should it be assumed that $\sigma_{e_1}{}^2 \leqq \sigma_{e_2}{}^2 \leqq \sigma_{e_3}{}^2$, then replacing $\sigma_{e_1}{}^2$ and $\sigma_{e_2}{}^2$ by $\sigma_{e_3}{}^2$ in the numerator and replacing $\sigma_{e_2}{}^2$ and $\sigma_{e_3}{}^2$ by $\sigma_{e_1}{}^2$ in the denominator above, one arrives at the liberal inequality for $\sigma_x{}^2$:

$$\sigma_x{}^2 < \frac{\sigma_{e_3}{}^4}{\sigma_{e_1}{}^2}\,.$$

But even relaxing the inequality thus and under the assumption, $\sigma_{e_1}{}^2 \leqq \sigma_{e_3}{}^2$, then the estimate of Method B_3 therefore will still not be as precise as that of Method A_3 unless

$$\sigma_x{}^2 < \sigma_{e_3}{}^2.$$

Generally speaking, however, it should be desirable for the precision of measurement or variance in errors of measurement (such as $\sigma_{e_1}{}^2$, $\sigma_{e_2}{}^2$, or $\sigma_{e_3}{}^2$) to be considerably smaller than the variance in the characteristic measured, $\sigma_x{}^2$. The appropriate relation between $\sigma_x{}^2$ and $\sigma_{e_j}{}^2$ depends upon the purpose for which the measurements are taken and the cost of insuring that $\sigma_{e_j}{}^2$ is sufficiently small relative to $\sigma_x{}^2$.

To obtain an estimate of product variability when three instruments are used to measure the characteristic x, the average covariance of the readings of I_1, I_2 and I_3 is computed. That is

$$\begin{aligned}(17)\quad \text{est. }(\sigma_x{}^2) &= \tfrac{1}{3}\{s_{x+e_1,x+e_2} + s_{x+e_1,x+e_3} + s_{x+e_2,x+e_3}\}\\ &= s_x{}^2 + \tfrac{2}{3}s_{xe_1} + \tfrac{2}{3}s_{xe_2} + \tfrac{2}{3}s_{xe_3} + \tfrac{1}{3}s_{e_1e_2} + \tfrac{1}{3}s_{e_1e_3}\\ &\quad + \tfrac{1}{3}s_{e_2e_3} \text{ symbolically}\end{aligned}$$

with variance therefore equal to

$$\begin{aligned}(18)\quad \frac{2}{n-1}\sigma_x{}^4 + \frac{1}{n-1}\{&\tfrac{4}{9}(\sigma_x{}^2\sigma_{e_1}{}^2 + \sigma_x{}^2\sigma_{e_2}{}^2 + \sigma_x{}^2\sigma_{e_3}{}^2)\\ &+ \tfrac{1}{9}(\sigma_{e_1}{}^2\sigma_{e_2}{}^2 + \sigma_{e_1}{}^2\sigma_{e_2}{}^2 + \sigma_{e_2}{}^2\sigma_{e_3}{}^2)\}.\end{aligned}$$

In case it may be desirable to list the average observed values of the characteristic measured, as in acceptance tests of products (see example at end of paper), then σ_x^2 may be estimated by the following method which involves variances of differences in readings of the three instruments. The mean reading of the three instruments on the *i*th item is symbolically

$$x_i + \tfrac{1}{3}e_{i1} + \tfrac{1}{3}e_{i2} + \tfrac{1}{3}e_{i3}$$

and the variance of such means is symbolically

$$s_x{}^2 + \tfrac{1}{9}(s_{e_1}{}^2 + s_{e_2}{}^2 + s_{e_3}{}^2) + \tfrac{2}{3}(s_{xe_1} + s_{xe_2} + s_{xe_3}) + \tfrac{2}{9}(s_{e_1e_2} + s_{e_1e_3} + s_{e_2e_3}).$$

Subtracting 1/18th the sum of variances of differences in readings of the instruments taken two at a time, one obtains the unbiased estimate

$$\text{(19)} \quad \text{est. } (\sigma_x{}^2) = s^2{}_{x+\frac{1}{3}e_1+\frac{1}{3}e_2+\frac{1}{3}e_3} - \tfrac{1}{18}\left\{s^2{}_{e_1-e_2} + s^2{}_{e_1-e_3} + s^2{}_{e_2-e_3}\right\}.$$

If it were desirable to work with sums of the readings of I_1, I_2 and I_3 of each item, (19) may, of course, be rewritten as

$$\text{(20)} \quad \text{est. } (\sigma_x{}^2) = \tfrac{1}{9}\left\{s^2{}_{3x+e_1+e_2+e_3} - \tfrac{1}{2}\left[s^2{}_{e_1-e_2} + s^2{}_{e_1-e_3} + s^2{}_{e_2-e_3}\right]\right\}.$$

It is remarked that (17), (19) and (20) are identical estimates of product variability.

It may be noted that the average measurements of the individual instruments I_1, I_2 and I_3 are

$$\bar{x} + \bar{e}_1, \qquad \bar{x} + \bar{e}_2, \qquad \bar{x} + \bar{e}_3$$

respectively and the differences in average errors of measurements of the instruments in pairs are

$$\bar{e}_1 - \bar{e}_2, \qquad \bar{e}_1 - \bar{e}_3, \qquad \bar{e}_2 - \bar{e}_3$$

which, of course, can be computed from the data giving rise to three equations and three unknowns

$$\begin{aligned} \bar{e}_1 - \bar{e}_2 &= c_1 \\ \bar{e}_1 - \bar{e}_3 &= c_2 \\ \bar{e}_2 - \bar{e}_3 &= c_3. \end{aligned}$$

However, the point of interest here is that although c_1, c_2 and c_3 are known from the data, the above simultaneous equations cannot be solved for the average errors of measurement $\bar{e}_1$, $\bar{e}_2$ and $\bar{e}_3$ since the equations are obviously linearly dependent. Of course, the size of $\bar{e}_1$, $\bar{e}_2$ and $\bar{e}_3$ may be determined relative to a standard.

7. THE SEPARATION OF s_x^2 AND $s_{e_j}^2$ WITH FOUR INSTRUMENTS

If four instruments I_1, I_2, I_3 and I_4 are used to make measurements on the same n items, the observed values or readings of the instruments can be represented as follows:

Measurements by I_1	Measurements by I_2	Measurements by I_3	Measurements by I_4
$x_1 + e_{11}$	$x_1 + e_{12}$	$x_1 + e_{13}$	$x_1 + e_{14}$
$x_2 + e_{21}$	$x_2 + e_{22}$	$x_2 + e_{23}$	$x_2 + e_{24}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n + e_{n1}$	$x_n + e_{n2}$	$x_n + e_{n3}$	$x_n + e_{n4}$

From the discussion so far, it is apparent that $\sigma_{e_1}^2$ and σ_x^2 may be estimated by considering the readings of I_1 with any one of the instruments I_2, I_3 and I_4—similarly for estimating $\sigma_{e_2}^2$, $\sigma_{e_3}^2$ and $\sigma_{e_4}^2$. Also, the variances in errors of measurement and variance in the characteristic measured may be estimated by considering the readings of any three of the four instruments at a time. These two problems have already been treated above. It remains, therefore, to consider simultaneously the results of measurement of all four instruments and to compare the precision of estimates thus obtained with that for the case involving three instruments.

Method A_4:

Using the variance of the readings of I_1 and the covariances of the readings of the four instruments taken two at a time, an estimate of $\sigma_{e_1}^2$ is given by

$$\begin{aligned}\text{est. }(\sigma_{e_1}^2) &= s^2_{x+e_1} - \tfrac{2}{3}(s_{x+e_1,x+e_2} + s_{x+e_1,x+e_3} + s_{x+e_1,x+e_4}) \\ &\quad + \tfrac{1}{3}(s_{x+e_2,x+e_3} + s_{x+e_2,x+e_4} + s_{x+e_3,x+e_4}).\end{aligned} \tag{21}$$

On the other hand, by forming the six columns of differences in readings of the four instruments taken two at a time, an estimate of $\sigma_{e_1}^2$ is given by

$$\text{est. }(\sigma_{e_1}^2) = \tfrac{1}{3}\left\{s^2_{e_1-e_2} + s^2_{e_1-e_3} + s^2_{e_1-e_4} - \tfrac{1}{2}\left[s^2_{e_2-e_3} + s^2_{e_2-e_4} + s^2_{e_3-e_4}\right]\right\}. \tag{22}$$

Estimates (21) and (22) are equivalent and symbolically equal to

$$s_{e_1}^2 - \tfrac{2}{3}(s_{e_1e_2} + s_{e_1e_3} + s_{e_1e_4}) + \tfrac{1}{3}(s_{e_2e_3} + s_{e_2e_4} + s_{e_3e_4}).$$

Consequently, the variance of estimate (21) or (22) is

$$(23)\quad \frac{2}{n-1}\sigma_{e_1}^4 + \frac{1}{n-1}\left\{\tfrac{4}{9}(\sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_1}^2\sigma_{e_4}^2) + \tfrac{1}{9}(\sigma_{e_2}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_4}^2 + \sigma_{e_3}^2\sigma_{e_4}^2)\right\}.$$

Estimates of $\sigma_{e_2}^2$, $\sigma_{e_3}^2$ and $\sigma_{e_4}^2$ by the above procedures and the variances of such estimates can be formed simply by rotating the subscripts e_1, e_2, e_3 and e_4 in (21) or (22) and (23) appropriately.

The estimate (21) or (22) of $\sigma_{e_1}^2$ and its variance (23) are entirely free of x, the characteristic measured, as is the estimate (8) or (9) employing three instruments. Estimate (21) or (22) using four instruments will be more precise than estimate (8) or (9) using three instruments provided

$$\tfrac{4}{9}\sigma_{e_1}^2\sigma_{e_2}^2 + \tfrac{4}{9}\sigma_{e_1}^2\sigma_{e_3}^2 + \tfrac{4}{9}\sigma_{e_1}^2\sigma_{e_4}^2 + \tfrac{1}{9}\sigma_{e_2}^2\sigma_{e_3}^2 + \tfrac{1}{9}\sigma_{e_2}^2\sigma_{e_4}^2 + \tfrac{1}{9}\sigma_{e_3}^2\sigma_{e_4}^2 < \sigma_{e_1}^2\sigma_{e_2}^2 + \sigma_{e_1}^2\sigma_{e_3}^2 + \sigma_{e_2}^2\sigma_{e_3}^2$$

i e provided

$$\sigma_{e_4}^2 < \frac{5\sigma_{e_1}^2\sigma_{e_2}^2 + 5\sigma_{e_1}^2\sigma_{e_3}^2 + 8\sigma_{e_2}^2\sigma_{e_3}^2}{4\sigma_{e_1}^2 + \sigma_{e_2}^2 + \sigma_{e_3}^2}$$

This condition appears easy to meet, since, for example, if $\sigma_{e_1}^2 = \sigma_2^2 = \sigma_{e_3}^2 = \sigma^2$, then one has only to be sure that

$$\sigma_{e_4}^2 < 3\sigma^2.$$

Moreover, if it is assumed that $\sigma_{e_1}^2 \leqq \sigma_{e_2}^2 \leqq \sigma_{e_3}^2$, then estimate (21) will surely be more precise than estimate (8) or (9) so long as

$$\sigma_{e_4}^2 < \frac{5\sigma_{e_1}^2\sigma_{e_1}^2 + 8\sigma_{e_1}^2\sigma_{e_1}^2 + 5\sigma_{e_1}^2\sigma_{e_1}^2}{4\sigma_{e_3}^2 + \sigma_{e_3}^2 + \sigma_{e_3}^2} \leq \frac{3\sigma_{e_1}^4}{\sigma_{e_3}^2}$$

or if

$$\sigma_{e_4}^2 < 3\sigma_{e_1}^2.$$

It is believed that this condition may be met easily in practice. Consequently, the use of an additional measuring instrument will apparently decrease the variances of the estimates of precision of measurement.

8. THE SEPARATION OF s_x^2 AND s_j^2 WITH N INSTRUMENTS

Defining e_{ij} as before to be the error in measurement of the ith item ($i = 1, 2, \cdots, n$) by the jth instrument, it is easy to extend the above results to any number, N, of instruments which are employed to measure the items, x_i. An estimate of the variance in errors of measurement of instrument I_1 is for $N \geqq 3$ given by

$$\begin{aligned}\text{est. } (\sigma_{e_1}^2) = s^2_{x+e_1} - \frac{2}{N-1}\left\{\sum_{r=2}^{N} s_{x+e_1,x+e_r}\right\} \\ + \frac{2}{(N-1)(N-2)}\left\{\sum_{2\leqq j<k}^{k=N} s_{x+e_j,x+e_k}\right\}\end{aligned} \tag{24}$$

which is symbolically equal to

$$s_{e_1}^2 - \frac{2}{N-1}\sum_{r=2}^{N} s_{e_1e_r} + \frac{2}{(N-1)(N-2)}\sum_{2\leqq j<k}^{k=N} s_{e_je_k}$$

so that the variance of (24) is given by

$$\begin{aligned}\frac{2}{n-1}\sigma_{e_1}^4 + \frac{1}{n-1}\left\{\frac{4}{(N-1)^2}\sum_{r=2}^{N}\sigma_{e_1}^2\sigma_{e_r}^2\right. \\ \left. + \frac{4}{(N-1)^2(N-2)^2}\sum_{2\leqq j<k}^{k=N}\sigma_{e_j}^2\sigma_{e_k}^2\right\}.\end{aligned} \tag{25}$$

Formulae for estimates of $\sigma_{e_2}^2$, $\sigma_{e_3}^2$, etc. and variances of such estimates may be found by obvious rotations of the subscripts.

The estimate of variability in the product is taken as the average of the covariances, i.e.

$$\begin{aligned}\text{est. } (\sigma_x^2) &= \frac{2}{N(N-1)}\sum_{1\leqq r<s}^{s=N} s_{x+e_r,x+e_s} \qquad (N \geqq 2) \\ &= s_x^2 + \frac{2}{N}\sum_{r=1}^{N} s_{xe_r} + \frac{2}{N(N-1)}\sum_{1\leqq r<s}^{s=N} s_{e_re_s} \text{ symbolically}\end{aligned} \tag{26}$$

so that the variance of estimate (26) is given by

$$\begin{aligned}\frac{2}{n-1}\sigma_x^4 + \frac{1}{n-1}\left\{\frac{4}{N^2}\sigma_x^2\sum_{r=1}^{N}\sigma_{e_r}^2\right. \\ \left. + \frac{4}{N^2(N-1)^2}\sum_{1\leqq r<s}^{s=N}\sigma_{e_r}^2\sigma_{e_s}^2\right\}.\end{aligned} \tag{27}$$

Should it be desirable to work with differences in errors of measurements of the instruments taken two at a time, an estimate of $\sigma_{e_1}^2$ is given by

$$(28)\quad \text{est.}\ (\sigma_{e_1}^2) = \frac{1}{N-1}\left\{\sum_{r=2}^{N} s^2_{e_1-e_r} - \frac{1}{N-2}\sum_{2\leq j<k}^{k=N} s^2_{e_j-e_k}\right\}. \quad (N \geq 3)$$

Estimates (24) and (28) are identical and hence have the same variance.

Using the average measurement of the N instruments on x_i and differences in errors of measurement, an estimate of product variability is

$$(29)\quad \text{est.}\ (\sigma_x^2) = s^2_{x+(1/N)(e_1+e_2+\cdots+e_N)} - \frac{1}{N^2(N-1)}\sum_{1\leq r<s}^{N} s^2_{e_r-e_s}.$$

On the other hand, using the sum of the measurements of the N instruments on x_i and columns of differences in errors of measurements, (29) may be rewritten as

$$(30)\quad \text{est.}\ (\sigma_x^2) = \frac{1}{N^2}\left\{s^2_{Nx+e_1+e_2+\cdots+e_n} - \frac{1}{N-1}\sum_{1\leq r<s}^{N} s^2_{e_r-e_s}\right\}.$$

It is easy to show that (29), (30) and (26) are identical estimates.

9. EXAMPLE ON ESTIMATING PRECISION OF MEASUREMENT AND PRODUCT VARIABILITY[6]

In Table I there are listed the individual burning times of powder train fuzes as measured by each of three observers on 30 rounds of ammunition which were fired from a gun. The fuzes were all set for a burning time of ten seconds.

The burning time of a fuze is defined as the interval of time which elapses from the instant the projectile leaves the gun muzzle until the fuze functions the projectile. The times given in Table I are measured by means of electric clocks. A switch on the gun muzzle starts three different electric clocks as the gun is fired and each observer stops his clock the instant he sees the flash or burst of an individual round. Each timer, of course, stops his clock independently of the other two timers. In Table I is given also the mean time of the three observers on individual rounds. The average time-to-burst of the 30 fuzes is taken as the average of the 30 mean times with the effect of reaction time (which is known accurately) subtracted therefrom. The problem here is to

---

[6] The data for this example are given also in Colonel L. E. Simon's paper, "On the Relation of Instrumentation to Quality Control," *Instruments*, vol. 19, November, 1946.

TABLE I

Fuze Burning Times (seconds)			Mean Times (seconds)
Observer A	Observer B	Observer C	
10.10	10.07	10.07	10.080
9.98	9.90	9.90	9.927
9.89	9.85	9.86	9.867
9.79	9.71	9.70	9.733
9.67	9.65	9.65	9.657
9.89	9.83	9.83	9.850
9.82	9.75	9.79	9.787
9.59	9.56	9.59	9.580
9.76	9.68	9.72	9.720
9.93	9.89	9.92	9.913
9.62	9.61	9.64	9.623
10.24	10.23	10.24	10.237
9.84	9.83	9.86	9.843
9.62	9.58	9.63	9.610
9.60	9.60	9.65	9.617
9.74	9.73	9.74	9.737
10.32	10.32	10.34	10.327
9.86	9.86	9.86	9.860
10.01	lost	10.03	10.020
9.65	9.64	9.65	9.647
9.50	9.49	9.50	9.497
9.56	9.56	9.55	9.557
9.54	9.53	9.54	9.537
9.89	9.89	9.88	9.887
9.53	9.52	9.51	9.520
9.52	9.52	9.53	9.523
9.44	9.43	9.45	9.440
9.67	9.67	9.67	9.670
9.77	9.76	9.78	9.770
9.86	9.84	9.86	9.853

determine whether the electric clock is a satisfactory instrument for measuring burning times, provided the electric clocks are properly calibrated with regard to average time so that the systematic error of measurement may be reconciled. This question is answered by comparing the variance in errors of measurement with the variance in burning times of the fuzes.

In Table II there are given the algebraic differences between times of the three observers A, B, and C on each round.

It is to be noted that any one of the differences listed is not influenced by the level of burning time of an individual fuze and represents the actual difference in errors of measurement of the two observers involved. Thus, the variance of each of the columns headed A–B, B–C, and A–C, gives the variability of the difference in errors of measurement of the two observers compared. Therefore, taking the variances

TABLE II

DIFFERENCE BETWEEN OBSERVERS' TIMES (SECONDS)

A–B	*B–C*	*A–C*
+.03	0	+.03
+.08	0	+.08
+.04	−.01	+.03
+.08	+.01	+.09
+.02	0	+.02
+.06	0	+.06
+.07	−.04	+.03
+.03	−.03	0
+.08	−.04	+.04
+.04	−.03	+.01
+.01	−.03	−.02
+.01	−.01	0
+.01	−.03	−.02
+.04	−.05	−.01
0	−.05	−.05
+.01	−.01	0
0	−.02	−.02
0	0	0
—	—	−.02
+.01	−.01	0
+.01	−.01	0
0	+.01	+.01
+.01	−.01	0
0	+.01	+.01
+.01	+.01	+.02
0	−.01	−.01
+.01	−.02	−.01
0	0	0
+.01	−.02	−.01
+.02	−.02	0

of the three columns in Table II, the following three equations are arrived at:

$$\text{(a)} \qquad \begin{aligned} s^2_{e_1-e_2} &= .0007030 \text{ sec.}^2 \\ s^2_{e_1-e_3} &= .0008878 \text{ sec.}^2 \\ s^2_{e_2-e_3} &= .0003108 \text{ sec.}^2 \end{aligned}$$

where e_1, e_2 and e_3 represent the errors of measurement of the observers A, B, and C respectively.

Using formulae similar to (9) of the paper for estimates of $\sigma_{e_1}^2$, $\sigma_{e_2}^2$, and $\sigma_{e_3}^2$ it is found that

$$\text{(b)} \qquad \begin{aligned} \sigma_{e_1} &= .0253 \text{ sec.} \\ \sigma_{e_2} &= .0079 \text{ sec.} \\ \sigma_{e_3} &= .0157 \text{ sec.} \end{aligned}$$

The above figures give a direct comparison of the ability or precision of each of the observers. It is seen that observer B is perhaps the best of the three since his precision of measurement is given by a standard deviation of only .0079 second, whereas observer A is the poorest timer of the three. As a matter of fact, one can look at the columns in Table II headed A–B and A–C and note that for the first 10 rounds or so observer A had a definite lag in stopping his clock as compared to the other two observers.

An estimate of σ_x^2 is obtained by using formula (19), i.e. by subtracting 1/18th of the sum of the variances of differences (a) from the variance of mean times in column four of Table I. Thus,

$$\begin{aligned} \text{est. } (\sigma_x^2) &= .046098 - \tfrac{1}{18}\{.0007030 + .0008878 + .0003108\} \\ &= .04599 \\ \text{or} \qquad \sigma_x &= .2145. \end{aligned}$$

Thus, it is seen that the variation in burning times, σ_x, is about 8.5 times the variance in errors, σ_{e_1}, for the poorest operator and about 12 times the average variability in errors of measurement of the three operators combined. Consequently, it is found that the combination of electric clocks and operators provides an adequate measuring instrument.

10. SIGNIFICANCE TESTS*

The problems of significance tests for comparing σ_x^2, $\sigma_{e_1}^2$, $\sigma_{e_2}^2$, etc. and judging whether $\epsilon_1-\epsilon_2$, $\epsilon_1-\epsilon_3$, $\cdots$, $\epsilon_2-\epsilon_3$, $\cdots$, $\epsilon_{N-1}-\epsilon_N$ differ significantly from zero need to be investigated, although certain rough or approximate tests may be evident to the reader. In this connection, it is remarked, for example, that for the case of three instruments if $\sigma_{e_1}^2$, $\sigma_{e_2}^2$ and $\sigma_{e_3}^2$ are of the same order of magnitude, then the approximate number of degrees of freedom for the est. $(\sigma_{e_1}^2)$, est. $(\sigma_{e_2}^2)$ or est. $(\sigma_{e_3}^2)$ is $\frac{2}{5}(n-1)$. However, the reader is warned that questions of dependence among the estimates need to be investigated since they may have an important bearing on significance tests. It is hoped that an investigation of these problems can be undertaken in the near future.**

---

* While this paper was being processed for publication, the author developed sequential probability ratio tests for the use involving two instruments.

** The author is indebted to Professor W. G. Cochran and referees of this paper for valuable suggestions in preparation of the manuscript.

LIFE COMPANY INVESTMENTS AND THE CAPITAL MARKETS*

FERGUS J. MCDIARMID
The Lincoln National Life Insurance Company

1. Savings made through life insurance are at this time the most common by far of all types of individual savings. Life insurance accumulations at the present time account for something around one-fourth of the total savings of individuals.

2. Savings made through life insurance are a very stable form of savings, varying less with business conditions than other forms of monetary savings.

3. At the present time, life insurance savings form by far the most important source of capital which seeks mainly a secure rate of return and safety of principal, rather than tax exemption or chance for capital appreciation. Traditionally they have been invested very largely in evidences of long term debt.

4. In the decade prior to the war, life company funds seemed to be outgrowing traditional investment outlets.

5. However, at the present time this trend seems to be in reverse, with investment outlets, even in traditional form, becoming rather plentiful. With declining savings, higher prices, and a huge real capital demand both here and abroad, we seem to have entered a period of real capital scarcity.

6. The main new outlet developed in recent years for life insurance funds has been direct investment in real estate, both housing and commercial.

7. The tendency for life insurance companies to invest largely in debt media may create a problem from the point of view of some industries which should be financed largely through equities.

GROWTH OF LIFE INSURANCE ASSETS

THE ASSETS of life insurance companies domiciled in the United States passed the $50 billion mark late in 1947, and they are increasing at a net rate in excess of $3 billion a year. This represents a doubling of the prewar rate of asset increase, a rate which still lags somewhat behind the increase in the national income over that period. Not only does this building up of life company assets create an investment problem of great magnitude, but it has an important influence on the capital markets in determining the types of securities through which capital can most readily be raised. Also the acceptability of an industry as a medium for the investment of life insurance funds may determine the relative ease with which it can raise capital in any form.

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

LIFE INSURANCE FUNDS IN RELATION TO TOTAL SAVINGS

Just how important is the flow of monetary saving through the life insurance companies, in relation to total saving by our people? The sum total of saving by individuals has been estimated by the Department of Commerce and the Securities and Exchange Commission. Each year's saving represents the sum of a number of items, the main ones being the increase in holdings of currency and bank deposits, the increase in insurance and pension reserves, the increase in security holdings, including U. S. Savings Bonds, the increase in the net equity in dwellings, and the decrease in consumer debt.

In the boom year 1929, the increase in life insurance reserves accounted for over 40 per cent of the total estimated net savings of individuals. During the decade of the thirties, the life insurance reserve increase slowed down some; but it constituted a rising proportion of total savings of individuals, averaging about three-fourths of the total estimate of these during that period. During the war years, the total of individual saving was greatly swollen; and life insurance accumulations, although rising, did decline as a proportion of total savings to less than 10 per cent of the total estimate in 1944. Since they reached their wartime peak in that year, net savings of individuals have declined very sharply, while life insurance assets have continued to grow. As a result, life insurance accounted for 22 per cent of the total saving flow in 1946. Because of the continued decline in the total of new saving by individuals, probably because of high living costs and the greater availability of durable consumer goods, life insurance funds probably will increase as a proportion of the total in 1947 and may very well account for at least one-fourth of total individual saving.

Certain features of the American economy tend to increase the relative importance of savings made through life insurance. It has become an increasingly urbanized economy and one of continuously greater specialization. To an increasing extent, the tools of production are financed and owned on a collective basis rather than through direct ownership by individuals. An expanding proportion of savings takes place through monetary channels rather than through the direct ownership of property.

Also, because of ascending income tax schedules on high incomes and the enlargement of the middle class group from below by higher real wage levels, a large and increasing proportion of total savings is now derived from people in the middle income brackets. This is the group which probably makes the largest proportional part of its savings through life insurance rather than through the direct ownership of securities and property, including equities in businesses.

There is still further evidence of the importance of life insurance as a means of saving. According to a study of consumer savings in 1946, published in the *Federal Reserve Bulletin* for August, 1947, the payment of life insurance premiums constituted by far the most common form of individual savings. Out of every 100 spending units in the country that year, 76 saved through the payment of life insurance premiums, 18 saved through payments on mortgages, and 13 by payments into retirement funds. These figures are not mutually exclusive. The point I wish to make is that, as a means of popular saving, life insurance led the parade by a very wide margin, the number of spending units saving through life insurance being not greatly exceeded by the number of such units saving through all other media combined.

NATURE OF LIFE INSURANCE SAVINGS

Savings made through life insurance have certain noteworthy characteristics. First of all, they are relatively stable—much more so than the sum total of individual savings combined. The reasons for this can be readily surmised. As long as a life insurance policy is kept in force, the accumulation of a reserve becomes an automatic affair; and, in general, people tend to keep their life insurance in force even if they have to forego temporarily other types of savings. In brief, the institution of life insurance has tended to make the accumulation of monetary savings more of a continuous and automatic process than it would be otherwise.

Another characteristic of life insurance savings is the relatively circumscribed scope for their investment. By law and tradition, the investment of these has been limited largely to senior securities, which means very largely evidences of debt. In other words, their investment has been restricted very largely to bonds and real estate mortgages.

There are a number of reasons for this state of affairs. In the first place, life insurance assets have been traditionally regarded as trust funds and their investment in the most senior and secure investment media has seemed proper.

Second, the mathematical calculation of life insurance reserves assumes a fixed but certain rate of return, and it has seemed proper that these reserves should be backed by assets which appear best able to assure this type of return.

Third, the above two considerations have been reinforced by law. The laws of most important states specify that life insurance funds, aside from certain limited proportions, should be invested in evidences of debt.

Fourth, regulations governing the valuation of life insurance assets tend to favor investment in bonds and mortgages. While well secured bonds and mortgages may be carried at cost or at amortized values, stocks, both preferred and common, must be carried at current market values. This tends to limit equity investments to a low percentage of assets because of the possible effect on the company's surplus of a serious decline in stock market prices. As a practical matter, therefore, life company investment in stocks rarely approaches the legally permitted limit, and seldom exceeds about 5 per cent of the assets of an individual company.

Because of the quite rigid manner in which the liabilities of a life insurance company are calculated, security of principal and reliability of rate of return are more important considerations than the chance for capital appreciation. Also, because of the manner in which life companies are taxed, tax exemption of income from securities is of very little value to them. In these respects, their position differs radically from that of wealthy individuals in high income tax brackets, to whom the chance to substitute a capital gains tax of 25 per cent for the regular income tax schedule is a prime consideration. The position of life companies with respect to tax exemption of securities is considerably different from that of the commercial banks, whose earnings are taxed on the same basis as those of corporations in general.

Because of these factors, life companies are mainly buyers of types of securities which have largely lost their appeal to individuals, namely, relatively high grade corporate bonds and mortgages. While there is some overlapping of interest between the life insurance companies and commercial banks, in general their requirements are different. Banks, to provide backing for their demand liabilities, invest largely in short term securities; while life companies, because of the steady growth in their assets and large and comparatively reliable cash inflow, do not in general feel a pressing need to sacrifice rate of return for liquidity. Also, tax exemption in a security has real value to a commercial bank, while in recent years the life companies have largely disposed of their tax exempt municipal bonds. However, because of sharply higher yields on municipal bonds, the life companies are again beginning to buy them on a small scale.

As a result of these factors and also because of the very large volume of funds at their disposal, the life companies have become the most important source of long term capital in this country of the type which is interested primarily in security of principal and sureness of rate of return. They have replaced the individual investor in this role since the last great period of capital expansion in this country—that of the twenties.

The relative importance of life insurance funds in the economy and the restriction, legal and otherwise, placed upon their investment, have given rise to a good deal of discussion and to some action, as we shall see later.

SUPPLY AND DEMAND FACTORS

For a long period in the history of this country, up to about a dozen years ago, the fact that life companies were largely restricted to investment in debt securities did not seem to create any serious problem. Up to the year 1930, at least, the growth in life company assets was approximately matched by that in long term private debt, consisting of corporate bonds and real estate mortgages. Up to that time, such private indebtedness had been the traditional channel of life insurance company investment, and in 1930 accounted for 84 per cent of the total invested assets of the companies.

But whereas in the decade of the twenties a $12 billion increase in life company assets was matched by a $30 billion increase in long term private debt, during the fifteen years beginning with 1931, life company assets increased $26 billion while private debt declined by $22 billion. Or, putting it a little differently, in 1930 life company assets equalled about one-fifth of the total long term private debt outstanding, while at the end of 1945 they were equal to nearly two-thirds of the total of such debt.

The point is that during this latter period, the traditional forms of investment media did not prove adequate to absorb the increase in life company funds. During this period, the life companies took over an increasingly large proportion of the total outstanding private debt. At the present time, for example, they own over one-half of the total bonds of the electric, telephone, and gas divisions of the public utility industry.

In the years just prior to the war, the life companies purchased a substantial volume of U. S. Government bonds, largely for want of other investment outlets; and during the war, their holdings of these bonds mounted to reach a total of nearly one-half of the total assets of the companies. During the war period, the yield on long term U. S.

Government bonds of 2½ per cent compared favorably with a yield of less than ¼ per cent more on high grade long term corporate bonds. This spread in yields has since about doubled, the yield basis on which long term high grade corporate bonds can be sold in quantity being at least ½ per cent higher than that on longest U. S. Government bond issues.

INVESTMENT IN REAL ESTATE

Probably the most noteworthy attempt to expand the investment field for life company funds was that into the direct ownership of real estate. Such real estate investments take two main forms—housing development and the ownership of business properties. The latter are usually leased to corporations of satisfactory credit standing on a basis to give the life company a moderate return and amortization of the investment over an extended period of years.

In recognition of the need for additional investment outlets, as well as the social desirability of permitting life company funds to flow directly into medium-cost housing (there is no low-cost housing any more), the laws of a considerable number of states have been amended to permit such investment. At the present time, the law of New York State, where life companies having about 60 per cent of the total assets of the industry are domiciled, permits investment of up to 10 per cent in housing projects built by the companies, and an additional 3 per cent in business or residential real estate. New Jersey, also an important state life insurance-wise, permits an investment of 5 per cent of assets in real estate held for business or residential purposes. Massachusetts currently permits 10 per cent of assets in rental housing and 10 per cent in other real estate. Connecticut and Wisconsin have catch-all provisions in their laws which permit investment in real estate or other investment media not otherwise authorized. Indiana has a law which permits investment in commercial real estate, subject to its lease to a corporation of defined credit standing, and in residential real estate.

THE PRESENT LARGE CAPITAL DEMAND

In the quite recent past, many people were fearful as to how far the lack of balance between supply and demand would go in depressing yields available for fixed interest securities of the type in which life companies had traditionally been invested. Today, for the time being at least, such fears are somewhat abated. A good many of us failed to foresee the huge postwar demand for capital and the effect of inflation in stimulating that demand. The rise in prices has had the effect of requiring from $1.50 to $2.00, or even more, to the work formerly done by $1.00.

Take the case of the utilities. From reading prospectuses, the impression is gained that many electric companies will make additions to their plants in the five years beginning with 1947 which will cost in dollars around 50 per cent of the cost of their existing plants. It costs about two-thirds more now than before the war to install a kilowatt of steam generating capacity.

American Telephone & Telegraph is currently engaged in a construction and financing program which will add about 40 per cent to its recent capitalization.

Many manufacturing and mercantile concerns have suddenly awakened to the fact that they need a lot of new money and need it rather badly—so badly that they are not in too good a position to argue about terms. It costs a lot more to carry inventory at present prices, and some inventories are unbalanced and a bit slow moving.

The recent demand for mortgage money, inflated by high real estate prices, has been particularly great; and some mortgages which were eagerly sought for a year ago now go abegging. Some life insurance companies have found themselves overly committed in their mortgage acquisitions and have practically withdrawn from that field.

Apparently, what we are witnessing is a sharp rise in the demand for investment capital, coupled with a continued decline in savings. High prices have probably reduced the ability of people to save. In any case, the interest on high grade long term bonds has increased by over ½ per cent in the last 1½ years, and the trend is upward. It currently appears that there will be a very large supply of such bonds sold for new money purposes over the next few years. Also, a yield of 3 per cent, which looked pretty bad to the institutional investor on the downside, now looks pretty good going up. With greener pastures available in their traditional investment fields, there is a greatly lessened incentive for life companies to seek new investment channels at this time.

FROM THE CAPITAL RAISER'S POINT OF VIEW

So far we have looked at this matter largely from the point of view of the life insurance investor. We will now attempt to look at it briefly from the point of view of the party requiring capital. Take the public utility industry as an example. It is perhaps the prime example of an industry requiring a large amount of funds of the type supplied by the life companies, that require a reasonably steady rate of return with capital appreciation as a quite secondary consideration.

It has come to be regarded as proper by the regulatory authorities that a utility should raise not much over half of its required capital in the form of debt and the balance by the sale of equity securities—preferred and common stocks—or by retained earnings. As long as they maintain their credit position, the utilities can probably raise capital by the sale of these bonds to life insurance companies. But what about balancing the sale of these bonds by that of equity securities on which, among other things, the maintenance of this credit position depends? During the last three years, the life companies have bought a rather large volume of utility and other preferred stocks. Recently, however, they have turned extremely cold toward this class of security, partly because some life companies have about reached their limit in this field, due to the reasons previously outlined, and partly because of the quite bad recent market behavior of these securities. Life companies have never been substantial buyers of utility common stocks.

Because of this situation it seems likely that the utilities are going to have trouble balancing their bond sales with stock sales unless they are prepared to do their stock financing on a relatively high cost basis. In addition, recent events have proved that they must also increase the attractiveness of their preferred stock financing by the insertion of sinking funds or convertibility features. They may even be forced to substitute convertible debentures for equity securities in their capital structures as far as the raising of new money is concerned. The substantial withdrawal of the life companies from the preferred stock market has made clear the extent to which they have rather recently come to dominate that market. Their dominance of the corporate bond market has been clear for at least a decade.

The sale of bonds has not been the traditional way for most branches of American industry to raise capital. In the past, only a few branches of our industry—the most important of which are steel, rubber, meat packing, and tobacco—have raised any appreciable part of their capital through the sale of bonds. On the other hand, some leading divisions of industry—including chemicals, electrical manufacturing, and many others—have been financed largely through equities. A good many types of industry, because of the relative irregularity of their earnings, would probably be well advised to avoid the use of contractual obligations for raising capital, if at all possible. However, the relative ease with which the funds of life companies can be tapped through the debt route, compared with the very great difficulty of tapping them through any other route, may put pressure on some industries to sell bonds when preferred or common stocks would be more suitable.

The emergence of the life companies as the main suppliers of fixed return capital in this country has had a major influence on the underwriting and distribution of securities. The selling of securities in large blocks to life companies, plus competitive bidding, have caused new issues to be handled on very small gross profit margins to underwriters, which would have been impossibly small if sale and distribution had to be made to a host of individual buyers. In view of the current attitude of the life companies toward preferred stocks, these may again have to be sold to individuals; the gross profit to underwriters at which they are handled will have to widen materially. The coupon rates also will have to be increased materially above recent levels to be attractive to individuals in substantial income tax brackets.

It is a well-known fact that keen rivalry exists among the large life insurance companies in the making of investments. They make their decisions on a highly individualistic basis and frequently pursue divergent investment policies. This, however, does not obviate the fact that under present conditions, success or failure in the underwriting of corporate bond issues of any size depends largely upon the willingness of at least some of the largest life companies to buy substantial amounts. Not only does this serve directly to get the bonds sold, but it no doubt influences the action of smaller investing institutions.

The dominance of the corporate bond market by the life companies has, of course, led to the private deal. It is probably natural that if a group of life companies are going to take most of an issue anyway, they should band together to shape the terms as much as possible to their liking, and as a secondary consideration save underwriters' profits. From the point of view of investing institutions, the advantages of such deals are that stronger protective terms can be imposed than in corresponding public issues, and a little better yield obtained.

From the point of view of the borrower, advantages include the avoidance of SEC registration, the chance to deal with only a limited number of professional investors in case a change in the terms is later desired, and the certainty of sale. Often an investment banker acts as agent in presenting the deal. In the case of a small issue of securities, particularly if put out by a borrower of secondary or unestablished credit standing, the saving to the borrower of going the private route may be very great indeed.

MORTALITY STATISTICS FOR LIFE INSURANCE UNDERWRITING*

EDWARD A. LEW
Metropolitan Life Insurance Company

The paper deals chiefly with those mortality statistics, compiled from the experience of insured persons, that have been used as a guide in life insurance underwriting. Properly interpreted, these statistics shed considerable light on the factors influencing human longevity. The use made of other mortality statistics in the underwriting of life insurance risks is also indicated.

LIFE insurance companies have conducted many investigations of mortality to obtain the statistics needed for the calculation of premiums, as a guide in the selection of risks, and for other purposes. Aside from serving their primary purposes, these studies contain a wealth of information regarding the factors influencing human longevity. In order to interpret the results intelligently, it is necessary to know something about the particular purposes of the various studies, the methods used in the compilation of the statistics, and their limitations.

This paper is confined to a consideration of the mortality statistics for life insurance underwriting. The term "life insurance underwriting" is applied to the procedures used by life insurance companies in passing upon applicants for life insurance. Broadly speaking, the objective of these procedures is to appraise applicants' prospects of longevity so as to classify them either into the various acceptable categories that have been established or the type which should be rejected, and to determine the appropriate premium rates for applicants in the acceptable categories.

The aims of this paper are:

(1) To sketch briefly how mortality studies for underwriting began in this country,

(2) To indicate the extent to which underwriting considerations have shaped both the form and content of the mortality studies for life insurance underwriting,

(3) To review the methods used in the more important investigations made by life insurance companies to obtain information for underwriting purposes,

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

(4) To consider the principal findings of these investigations in the light of the limitations of the data, and

(5) To examine the usefulness for underwriting purposes of mortality statistics from sources other than the experience of insured lives.

ORIGIN OF MORTALITY STUDIES FOR UNDERWRITING

The efficacy of life insurance underwriting depends largely on knowledge regarding the effects on mortality of those characteristics of life insurance applicants which can readily be ascertained in practice. That detailed quantitative knowledge is not absolutely necessary is indicated by the earlier successful record of the life insurance business, first in Great Britain and later on this continent. Even in the first part of the eighteenth century, there was sufficient qualitative understanding of many factors affecting human longevity, based on common sense observations, to have led some English companies to charge extra premiums to travelers going to Ireland, to women of child-bearing age, and to persons who had not had smallpox. Generally speaking, however, life insurance underwriting before the advent of mortality statistics was primarily concerned with eliminating applicants who were in poor health or otherwise undesirable risks on account of intemperate habits, questionable morals, etc.

More exact knowledge of mortality, based on statistical investigations, had to wait until sufficient data had been accumulated and suitable methods for extracting the information devised. That the data became available as soon as they did is in no small measure due to the remarkable growth of life insurance in the United States. With increasing numbers of persons rejected for life insurance during the last decades of the nineteenth century, there arose in the course of time an effective demand for the insurance of substandard risks. American companies sought to make such insurance available and turned to their actuaries to devise methods and to determine premium rates for insuring substandard risks. In 1889 the Actuarial Society of America was organized and this body naturally brought to bear on these problems the techniques originated and the experience with mortality investigations gained by actuaries in Great Britain. No less opportune was the introduction at about this time of punch card equipment for making mortality investigations. As a result of these developments, and because of the accumulation of a sizeable volume of data by reason of the rapid growth of life insurance, the Actuarial Society of America was able by 1901 to undertake an extensive intercompany investigation (Specialized Mortality Investigation) of various factors affecting the mortality of insured lives. This investigation was the forerunner of many similar efforts, both intercompany and by individual companies, designed to provide mortality statistics for life insurance underwriting.

MORTALITY STATISTICS ADAPTED TO UNDERWRITING CONSIDERATIONS

Mortality statistics derived from the experience among insured lives furnish the most pertinent guide for life insurance underwriting. This is because they show the actual results obtained for the type of information normally elicited in life insurance applications and from other sources available to life insurance companies. While our general knowledge of the factors influencing mortality has been greatly enriched by the increasing numbers of statistical studies made by research groups, hospitals, physicians, and others, such studies have sometimes only limited applicability to life insurance underwriting.

Mortality factors considered in underwriting. Most of the mortality investigations made by life insurance companies to obtain information for underwriting purposes have been concerned with those particular characteristics known to affect mortality about which information could readily be obtained by an insurance company. In addition to age, the principal characteristics so used in appraising applicants' prospects of longevity have been: (1) occupation, (2) height and weight, (3) physical impairments brought to light by the usual type of medical examination for life insurance, (4) personal medical history and family history as disclosed in response to questions asked in life insurance applications, and (5) habits as reported by independent inspections. These are by no means the only characteristics highly correlated with mortality variations or even necessarily the most important ones, but over the years it has been found possible to secure satisfactory information about them through the channels available to life insurance companies. Details of these characteristics are obtainable more accurately on applications for Ordinary insurance than on other forms of life insurance and as a result studies of mortality made for underwriting purposes have been predominantly based on the experience of lives insured under Ordinary policies.

Numerical rating system. Gradations in the factors commonly considered in underwriting are expressed in terms of the excess mortality they are expected to bring about among lives otherwise acceptable for insurance at standard premium rates. This excess mortality for a group having a specific characteristic, stated as a percentage of the mortality among a similarly constituted group of lives acceptable for standard insurance, is referred to as the rating for the characteristic; by adding the ratings for various characteristics of a risk, treating interrelated characteristics as a single factor, a numerical appraisal of the total risk is obtained. This method of appraising longevity in numerical terms (known as the numerical rating system) requires that the ratings for the various characteristics be based on actual experience whenever possible.

Mortality ratios. To facilitate interpretation of mortality statistics as ratings, mortality data for insured lives have customarily been summarized as ratios of the deaths actually experienced in a group to the deaths that would have occurred had the group been composed of otherwise comparable lives acceptable at standard premium rates. Such ratios show only the *relative mortality* associated with a particular characteristic during a given period of observation, but this is the criterion usually employed in classifying risks for underwriting purposes. It should be noted that in determining both the numerator and the denominator of these ratios, the number of policies terminated by death (or the amounts of claims paid) has generally been used in lieu of the number of persons involved. While the mortality ratios by number of policies (or amounts of insurance) may occasionally exhibit extraneous fluctuations due to a large number of policies (or a substantial amount of insurance) being involved in a single death, they have the advantage of reflecting more accurately the financial effect of the extra mortality traceable to particular characteristics.

Risks appraised as members of a class rather than as individuals. A very precise appraisal of various characteristics in numerical terms is neither practicable nor necessary for classifying applicants as life insurance risks. Only on very large amounts of insurance does it become worth while to make elaborate investigations. For most practical purposes, it is sufficient to be able to make decisions as to the broad class of risk to which a particular applicant should be assigned. The averaging of risks is a fundamental principle of insurance and the underwriter accordingly looks upon risks more nearly as members of a class than as individuals. In line with this principle, ratings are more appropriately regarded as mean values of a reasonable range of variation. In making mortality studies of various characteristics for the purpose of determining ratings, actuaries have, therefore, often preferred to deal with risks in broad groups, whose heterogeneity corresponded to that of the rating classifications, rather than subdivide the experience into more homogeneous but smaller groupings subject to large chance fluctuations.

Risks classified on basis of characteristics at time of application. Life insurance policies do not as a rule contain any provisions which enable the company to change the classification of a risk after issue, even though the risk increased materially. Applicants for life insurance must, therefore, be appraised on the basis of their characteristics at time of application for the insurance, taking into consideration the possibility of the risk's increasing. This has important implications as to the type of mortality statistics needed for life insurance underwriting. For characteristics that can be changed by the insured at his own volition, such as occupation or aviation activities, the ratings must be determined from an experience of closed groups of individuals having the particular characteristics at time of entry; such experience should reflect the effect of the usual proportion of individuals changing to a more hazardous status as well as the proportion of individuals who change to a less hazardous status and leave the group. For characteristics where the excess mortality does not manifest itself until many years have elapsed, such as overweight or moderate elevation in blood pressure, the ratings must be determined from experience of sufficiently long duration.

Incidence of mortality by age and duration. Because premiums for life insurance depend to a marked degree on the incidence of mortality by age and duration from issue, mortality statistics for life insurance underwriting have usually been compiled so as to bring out the variations in mortality by age and duration. Such information is not only helpful in appraising and classifying risks but indicates also the type of extra premium to be charged—constant with age, increasing with age, decreasing with duration, etc. Mortality statistics by number of policies and especially by amounts of insurance are, of course, more suitable than mortality statistics by lives for the purpose of determining extra premium rates.

METHODS USED IN MORTALITY INVESTIGATIONS OF INSURED LIVES

Intercompany mortality investigations. The principal investigations made to obtain mortality statistics for underwriting purposes have been carried out by the Joint Committee on Mortality of the Actuarial Society of America and the Association of Life Insurance Medical Directors. The first study made by this Committee (Medico-Actuarial Mortality Investigation) was based upon the combined experience of a number of companies under Ordinary policies issued during the years 1885–1908 and included average heights and weights, the effects of build on mortality, the mortality of women and certain races, mortality by occupation and mortality for various medical impairments. Since that time, nine additional mortality investigations have been made by the Joint Committee, covering chiefly medical impairments and occupations.[1]

These later investigations have all dealt with the combined experience among lives insured under Ordinary policies in the larger life insurance companies of the United States and Canada. The usual procedure has been for individual companies to send to the Joint Committee summary sheets for each class of risk included in the study showing (1) the "exposed to risk," (2) the "expected" deaths, and (3) the actual deaths, tabulated by broad age groups at issue and appropriate groupings by duration from issue. As the basis for calculating the "expected" deaths in each investigation, the Joint Committee has constructed special mortality tables (often referred to as "basic tables") in "select" form, i.e., showing mortality rates by age groups at issue and duration from issue. These tables represent the companies' combined experience under Ordinary life insurance issued at standard premium rates for the period of years covered by the investigation. Special tables have also been prepared to serve as yardsticks of the expected mortality from the major causes of death.

The data for these mortality investigations have been assembled by the individual companies on punch cards, prepared in the larger companies for a sample of the Ordinary policies issued at standard premium rates and for all Ordinary policies issued at substandard premium rates. The punch cards have been coded for occupation, medical impairments, and certain other characteristics in accordance with codes specially prepared by the Joint Committee on Mortality. Occasionally, it has been necessary to go back to application papers in order to study characteristics which were not distinguished in the codes published by the Joint Committee.

From time to time, the Joint Committee has asked the companies for suggestions as to the classes of risks whose mortality it would be advisable to investigate. The classes favored have then been reviewed to determine whether the volume of data was large enough to warrant their inclusion in the study. For this purpose the larger companies have maintained up-to-date tabulations of the number of deaths on policies with rateable occupations, medical impairments, etc.

As a general rule, in selecting cases with a particular characteristic for inclusion in a study, it has been customary to eliminate those that also presented other characteristics which by themselves would have precluded the issuance of a policy at standard rates. On the other hand, cases accepted at standard rates have been included even though they might exhibit minor characteristics not of sufficient importance to require an extra premium but nevertheless of possible significance. Under these circumstances, the experience of a class may to a degree reflect the excess mortality attributable to the cases with minor characteristics. The proportion of cases with such minor characteristics has usually been a factor in determining whether a particular class should be studied, and in some investigations the cases with minor characteristics have been eliminated to obtain an experience more truly representative of the characteristic under study.

---

[1] Blood Pressure Study 1925, Joint Occupation Study 1928, Medical Impairment Study 1929; Supplement to Medical Impairment Study 1932, Impairment Study 1936; Occupational Study 1937; Impairment Study 1938; Blood Pressure Study 1939; Supplement to Blood Pressure Study 1939.

Shortcomings of intercompany studies. There has been some question as to whether mortality investigations of the combined experience of many companies with different underwriting practices and standards yield significant results for certain classes of risks. The point has been made that the mortality experience of companies having similar underwriting practices and standards could be interpreted more accurately and that the best guide for an individual company was its own experience. Because of different application forms, different medical examination requirements, and different attitudes toward indefinite reports, it has been found that the cases included by different companies in classifications which are not susceptible to a precise definition may differ markedly. However, the data included in classifications that can be clearly defined have been found to give a satisfactory degree of homogeneity. Present opinion favors a continuation of mortality investigations of the combined experience of many companies for classes where the data are homogeneous or where a sufficiently large experience could not be obtained otherwise.

Individual company and other mortality studies of insured lives. Studies of the experience of individual companies have also contributed much information regarding effects of various characteristics on mortality.[2] The results have generally been easier to interpret than those of intercompany investigations because of greater homogeneity of the data and the greater ease with which questionable points could be reviewed. On the other hand, small numbers have been more frequently encountered in individual company studies, so that recourse has sometimes been had to more advanced statistical techniques to gauge the reliability of the results. While individual companies' mortality investigations have as a rule been made from the same punch cards as those used in intercompany investigations, the methods employed have, of course, varied considerably. For instance, some companies have occasionally determined their experience for various classes of risks by relating all the deaths in a class to exposures built up from a sample of the policies issued.[3]

---

[2] Among the more notable of these have been a series of papers by A. Hunter and O. Rogers (*Transactions* of the Actuarial Society, Vols. XXIII to XXX) and another series by L. I. Dublin in collaboration with various others (*Proceedings* of the Association of Life Insurance Medical Directors, Vols. XV, XIX to XXVI, and XXVIII).

The techniques used in follow-up studies of individuals subjected to special examinations and of rejected risks also deserve attention because of their applicability to other fields. The usual practice has been to trace individuals chiefly by reference to their insurance holdings in various companies. It now appears possible that Social Security records might be utilized for this purpose, if suitable arrangements can be worked out with the Social Security Administration, and the records are handled in such a manner that particular individuals cannot be identified.

FINDINGS OF THE PRINCIPAL MORTALITY STUDIES OF INSURED LIVES

In looking at the findings of the principal mortality studies of insured lives, it is necessary to keep in mind that they represent the experience among persons insured under Ordinary policies. Such persons comprise a highly selected portion of the general population; they are predominantly white male adults, and in the past at least were drawn chiefly from the more prosperous classes of the population. Even when insured at substandard premium rates, they usually do not present characteristics that result in a drastic reduction in longevity.

While absolute death rates for various classes of risks can be derived from the mortality ratios developed in these investigations, the mortality ratios themselves indicate only the relative mortality of different classes of risks, compared with the level of mortality among lives insured at standard premium rates during the same period of years. Since the level of mortality among lives insured at standard rates has been changing over the years, different yardsticks of mortality have in effect been used in different studies. As the level of standard mortality has declined, the relative mortality of risks whose absolute mortality rates remained unchanged has increased, while the relative mortality of risks whose absolute mortality rates decreased more rapidly than that of standard lives has declined.

---

[3] See paper "Sampling Errors in Life Insurance Mortality and Other Statistics" by Donald Cody to be published in the September 1948 issue of the *Journal of the American Statistical Association*.

Most of the mortality investigations of insured lives represent experiences spanning relatively short periods of time. The great volume of the experience has usually been in the years immediately following issue of the insurance, while the experience at the longer durations has sometimes been too scant to give reliable information. This is a serious shortcoming in the case of those classes of risks where the extra mortality is deferred, since the results will tend to understate the excess mortality that may actually be experienced over the entire period of insurance.

Despite these and other limitations, mortality studies of insured lives provide much pertinent information of a broad qualitative nature as to the effects on mortality of occupations, height and weight, physical impairments, and medical histories. In interpreting the experience of particular groups of lives, it is essential to have a clear idea as to the type of cases included in each classification. These points are illustrated in the following discussion of the findings of the more important studies.

Occupations. The first extensive intercompany investigation (Specialized Mortality Investigation) covered the experience during the years 1870–1900. It included a study of 35 groups involving an occupational hazard, among them commissioned officers, the more important hazardous trades, liquor distribution, and railroad service.

The Medico-Actuarial Mortality Investigation, which covered the years 1885–1909, included 168 groups of occupations. The most important of these were in the liquor industry, railroad service, metal trades, and mining. This study showed that mortality was particularly high among underground miners, railroad engineers and firemen, structural iron workers, stone cutters, and many occupations in the liquor industry. In certain hazardous occupations, the unusually low mortality experienced was believed to be due to very careful selection of risks by the companies and to the short average duration of the experience.

The most comprehensive investigation of occupations thus far conducted was the Joint Occupation Study. It called for the study of some 430 groups of occupations and covered the years 1920–27 for occupations with a major accident hazard and the period 1915–27 for other occupations. The occupations studied ran practically the entire gamut of occupational hazards commonly met with among applicants for Ordinary life insurance. Very high mortality was found among metal miners, coal miners, oil and gas field rig builders, linemen in electric light and power plants, chippers of metal, laborers in iron and steel works and foundries, and bartenders. Tuberculosis was found to be high in some 25 occupational classes, chiefly among unskilled labor and in occupations characteristic of the lower income groups, and also among those exposed to the silica dust hazard, notably stone cutters. A special study was also made of the accidental death rates, differentiating between those due to occupational and non-occupational accidents. The highest accidental death rates were found among electric light and power linemen, oil and gas field rig builders, underground miners, structural iron workers and freight train brakemen.

The most recent of the investigations dealing with occupational hazards—the Occupational Study of 1937—covered the years 1925–36 and included 87 occupational groups. Among them were many for which the results in the Joint Occupation Study had been inconclusive. Very high mortality was found among metal miners (high accidental death rate, tuberculosis and pneumonia), motorcycle police (high accidental death rate), anthracite coal miners (high accidental death rate, tuberculosis and pneumonia), chippers of metal (high tuberculosis), electric light and power linemen (high accidental death rate), and stone cutters (high tuberculosis). The study showed a decrease in the mortality of laborers in non-hazardous industries, believed to reflect in some measure the improvement in the living conditions of the lower income groups.

A review of the findings of these investigations indicates that the mortality in occupations with pronounced accidental death or health hazards has in general increased relative to the mortality of lives insured under Ordinary policies at standard premium rates; on the other hand, the relative mortality in many occupations where the excess mortality was at one time largely due to low living standards appears to have decreased. This supports the significant conclusion reached from various studies that the effects of occupation on mortality are, on the whole, more indirect than direct; mortality in most occupations is influenced more by the environment and the living standards associated with the occupation than by the accident and health hazards traceable to particular occupational duties.

Height and Weight. The first table of average weights according to height and age, based on a considerable volume of statistics, was prepared from life insurance data by the Association of Life Insurance Medical Directors in 1897 and was known as the Shepherd table.

The Medico-Actuarial Mortality Investigation developed new tables of average heights and weights by age for men and women and also included an extensive investigation of the effects of build on mortality. The height and weight tables were based on the records for 221,819 males and for 136,504 females, representing a random sample of the heights and weights, in street clothing (without overcoats), of persons to whom standard Ordinary policies were issued mainly during the years 1885 to 1900, inclusive. These tables of average heights and weights remain in general use today, because subsequent studies showed that there has been little change in average heights and weights over the years. The investigation of the effects of build on mortality brought out the following main conclusions:

(1) That mortality increases with decreasing weight at the younger ages,
(2) That the effect of underweight on mortality decreases with advancing age,
(3) That overweight is not a serious impairment at the younger ages but its adverse effect on mortality increases with advancing age,
(4) That the lowest mortality at ages under 35 is found among those 15 to 20 pounds above average but at ages 35–49 it occurs among those 5 to 10 pounds underweight,
(5) That young underweights have high mortality from tuberculosis and pneumonia,
(6) That overweights have high death rates from diabetes, nephritis, heart disease, and cerebral hemorrhage.

Average heights and weights were determined again in the Medical Impairment Study for persons to whom standard Ordinary policies were issued during the years 1909–27, but no new tables were published for ages 25 and over because the figures developed were found to be remarkably close to those established by the Medico-Actuarial Mortality Investigation. New tables of average heights and weights were, however, compiled for ages 10 to 24. The Medical Impairment Study also confirmed the major findings of the Medico-Actuarial Mortality Investigation in regard to the effect of build on mortality.

It appears likely that the relative mortality of both overweights and underweights has changed somewhat since the publication of the 1929 Medical Impairment Study. This study did not extend over the recent period of years during which mortality from tuberculosis and pneumonia showed a marked decline. Since these causes of death have in the past accounted for practically all of the excess mortality among underweights, it would be reasonable to suppose that the mortality of underweights has improved in relation to that of lives insured under Ordinary policies at standard premium rates. On the other hand, there has apparently been little change in the mortality from heart disease, diabetes, and other degenerative diseases, which have accounted for most of the excess mortality among overweights. This suggests that the relative mortality of overweights may have increased.

Physical Impairments and Medical Histories. The particulars given in life insurance applications concerning occupation and height and weight can as a rule be taken without qualification. However, this is not the case in much of the information elicited regarding physical impairments and medical histories.

The medical examination for life insurance consists usually of two parts. The first is made up of the applicant's own statements regarding details of his physical condition, history of illnesses, injuries, and medical treatment received, as well as some miscellaneous matters such as family history. The second part is in the nature of a report by the medical examiner on the physical examination of the applicant; this examination customarily embraces a check on the important organs of the body, height and weight, a blood pressure reading and urinalysis. Special tests such as an electrocardiogram, X-rays, and blood sugar tests are expensive and can be used only on large amounts of insurance or in other cases where they are clearly of great value. There are also practical limits on the extent to which indefinite statements made by the applicant or the medical examiner can be followed up. The result is that many of the descriptions of physical impairments and medical histories on life insurance applications represent relatively undifferentiated conditions. The findings of the physical examination itself cannot usually be compared in point of comprehensiveness and precision with those made by clinicians who are in position to study a patient for a longer period of time, more intensively and with his full cooperation. As a consequence, the mortality experienced among insured lives with certain physical impairments or medical histories furnishes a guide of only limited relevance for cases where specific and accurate diagnoses have been established.

Because persons suffering from serious organic diseases or from acute illnesses would not ordinarily apply for or be accepted for life insurance, the mortality studies of insured lives have perforce dealt with the less severe and non-acute conditions. With few exceptions, clinical medicine has not been able to follow large numbers of patients with relatively minor impairments for a sufficiently long period of time to determine their effect on mortality. Despite their limitations, mortality studies of insured lives constitute, therefore, a major source of information regarding the significance of many such impairments. The most valuable contributions to medical science made by mortality studies of insured lives have probably been with respect to factual conditions that can be appraised in numerical terms, such as build, blood pressure, urinalysis findings and pulse rate.

The earliest of the intercompany studies—the Specialized Mortality Investigation—included 24 groups of physical impairments and medical histories. However, only cases accepted at standard premium rates were included and some of the classes were described in terms no longer current. The Medico-Actuarial Mortality Investigation included 83 groups of physical impairments and medical histories, all but two of which were limited to cases accepted at standard premium rates. The list of impairments studied was not greatly different from that of conditions considered significant today. Among the many results of interest, it was found that persons with a rapid pulse rate of between 90 and 100 per minute showed substantially higher than normal mortality, while those with a pulse rate over 100 per minute experienced even higher mortality. Later studies corroborated these findings.

The most comprehensive of the intercompany investigations of physical impairments and medical histories was the Medical Impairment Study. It covered the experience on policies issued during the years 1909–27 and included 122 groups of impairments and histories, many being further subclassified by duration and severity. Cases accepted at extra premium rates as well as those accepted at standard premium rates were studied. Because of the inclusion of cases accepted at extra premium rates, the mortality associated with a number of more serious conditions was developed for the first time in an intercompany study.

Probably most important among the serious conditions were several classes of heart impairments, diagnosed primarily by reference to types of heart murmurs they produce. The mortality experienced on certain of the heart murmurs confirmed clinical opinion to the effect that they were relatively unimportant, but other types of murmurs were found to be associated with very high mortality. A heart murmur in conjunction with a moderate degree of hypertrophy of the heart and particularly with a history of rheumatic fever was shown to be a very serious condition. Heavy manual work was found to have an adverse effect on the mortality of those with heart disease.

The Medical Impairment Study included also the experience for other combinations of impairments such as heart murmurs with elevated blood pressure, albumin with casts, and albumin with elevated blood pressure. It was found that the combination of albumin and overweight presented a very much more serious impairment than the sum of the extra mortality associated with each of these conditions would indicate.

The Medical Impairment Study brought out clearly for the first time that even small elevations in systolic blood pressure are associated with a marked increase in mortality. This finding was at variance with general medical opinion and a number of additional investigations were made to learn more about the effects on mortality of variations in blood pressure. The most important of these was the 1939 Blood Pressure Study, which considered the variations in both systolic and diastolic pressures. The main conclusions of the several studies of blood pressure were as follows:

(1) A systolic blood pressure of 20 mm. above average or a diastolic blood pressure of 15 mm. above average results in distinctly higher mortality than normal.
(2) Mortality is heavier where both systolic and diastolic pressures are high.
(3) The combination of elevated blood pressure with certain minor impairments, particularly those of the cardiovascular-renal system, produces appreciably higher mortality than the sum of the extra mortality associated with the respective conditions.
(4) The lowest mortality is found in the group with systolic pressures from 5 to 15 mm. below average.

Two additional mortality studies supplemental to the Medical Impairment Study were made in 1936 and 1938. They covered 42 classes of physical impairments and medical histories. Further evidence was adduced by these studies as to the adverse effect on mortality of gastric and duodenal ulcers, removal and drainage of gall bladder, toxic goitre, syphilis, renal colic, irregular and intermittent pulse and albuminuria. The findings for several of these classes have probably lost some of their significance because of the recent advances in medical treatment and in surgery; the relative mortality of conditions that may require surgery, such as gastric and duodenal ulcers, is likely to have declined somewhat and the same probably applies to the mortality associated with syphilis.

STATISTICS FROM SOURCES OTHER THAN THE EXPERIENCE OF INSURED LIVES

It usually takes a long time to accumulate sufficient data for a mortality investigation of insured lives with relatively rare characteristics. For many types of risks it is virtually impossible to obtain mortality statistics based on the experience of insured lives. In such cases the underwriter naturally falls back on whatever other information is available. Another difficulty is that by the time the mortality statistics have been compiled, the relative hazard of a characteristic may have changed materially. This has occurred in the case of certain occupations where the working and living conditions have improved markedly in recent times, probably also to some extent for overweights and underweights because of the changing mortality trends, and for a few medical conditions where new and radical methods of treatment have been discovered. Since the underwriter is primarily concerned with the mortality likely to be experienced on various types of risks in the future, he must keep in close touch with changing conditions in industry, with developments in medicine and public health, and with mortality trends in general. The availability of more and better public health statistics, medical statistics, and general population statistics has been of great help to actuaries and underwriters in making judgments regarding the probable future mortality of different classes of risks.

Life insurance companies have made considerable use of the records available in governmental agencies and in industry to estimate death rates for many occupations. The statistics maintained by governmental agencies have made it possible to determine the mortality rates among personnel of the armed forces, scheduled airline pilots and state police, to mention only a few. It has been more difficult to derive death rates for various occupations in industry, because records for individual occupations are rarely segregated and because often only the deaths due to occupational accidents can be traced. Even this limited information, however, has been of value in considering occupations where accidents are the major hazard. It might be noted that appraisals of occupational hazards based on government or industry statistics represent usually the death rates among those actively at work. For underwriting purposes, it may be necessary to make allowances for the higher mortality among those in the occupation who have become disabled and left work and for the proportion of the individuals who change to more hazardous occupations in the normal course of events.

Statistics compiled by hospitals, physicians and public health agencies have also been widely consulted for life insurance underwriting. Clinical statistics have been helpful in determining appropriate ratings for the many relatively rare conditions on which life insurance experience can shed no light. Such data have also proved of value in indicating under what circumstances a relatively serious impairment, such as diabetes, previously considered uninsurable, may be accepted at extra premium rates. Although public health and population statistics have served chiefly to bring out changes in mortality trends, they too have occasionally been used as a basis for ratings. The application of clinical data to medical underwriting has presented many difficulties: for instance, the problem of translating the findings of accurate clinical diagnoses into the less definite categories found on life applications, and the difficulty of allowing for the difference between the details normally disclosed by a patient to his physician and the details stated on an application for life insurance, which may refer to conditions that have never received medical attention. It has also been necessary to make allowance for the relatively short term outlook in clinical medicine in contrast to the long term view of life insurance underwriting. The increasing use of follow-up studies by research groups, hospitals, and physicians indicates a growing appreciation of the need to study the long range effects on mortality of various medical conditions.

BIBLIOGRAPHY

Actuarial Society of America, *Specialized Mortality Investigation*, "Experience of Thirty-Four Life Companies upon Ninety-Eight Special Classes of Risks, 1903."

Bonnett, E. C. and E. A. Lew, "A Mortality Study of Systolic Heart Murmurs," *Proceedings* of the Association of Life Insurance Medical Directors of America, vol. 29, 1942, page 8.

Hunter, A., "Comments on the Medical Impairment Study (1929)," *Transactions* of the Actuarial Society of America, vol. 32, 1931, page 414.

Joint Committee on Mortality of the Actuarial Society of America and the Association of Life Insurance Medical Directors.
Medico-Actuarial Mortality Investigation, 1912.
1928 Joint Occupation Study, 1929.
1929 Medical Impairment Study, 1931.
Supplement to Medical Impairment Study, 1932.
1936 Impairment Study, 1936.
1937 Occupational Study, 1938.
1938 Impairment Study, 1938.
1939 Blood Pressure Study, 1940.
Supplement to Blood Pressure Study, 1941.

Mead, Franklin B., "The Evolution of Medical Selection and of Life Insurance Underwriting," *Record* of the American Institute of Actuaries, vol. 9, 1920, p. 1.

Shepherd, P., "Analysis of Blood Pressure Mortality," *Proceedings* of the American Life Convention (1941 Medical Section), page 61.

STATISTICAL PROBLEMS IN DEVELOPING VOLUNTARY MEDICAL EXPENSE INSURANCE PLANS*

MORTON D. MILLER
The Equitable Life Assurance Society of the United States

Of the 45,000,000 persons having hospitalization insurance, 20,000,000 have become covered by some form of insurance for the expenses of medical care. Most of this coverage has been provided since 1938, principally by the life insurance companies through group insurance. This paper discusses the statistical problems accompanying the development of medical expense insurance plans: the sources of data, the difficulties encountered in adapting the data to the purpose, and the nature and purpose of the statistics to be collected as experience unfolds. Insurance company experience with medical expense insurance is producing much valuable information regarding the demand for various medical services, the cost of such services and the administrative problems involved in providing for medical services on a prepayment basis.

THE TREMENDOUS growth of hospitalization insurance is a fairly familiar story. In the relatively short period since 1930, the number of persons insured has risen from almost nothing to the point where today more than 45,000,000 persons are covered for the cost of hospitalization. The development and spread of voluntary plans to insure the cost of doctors' services promises to be equally phenomenal.

It is only since 1938 that the established insurance companies have been insuring surgeons' charges for the performance of surgical operations and obstetrical procedures, and since 1943 that the companies have been covering the cost of other types of medical services. Now in 1947, over 20,000,000 persons are insured for surgical benefits, of whom more than ⅓ are covered in addition by some form of medical expense insurance for non-surgical services.

While community organizations similar to and sometimes affiliated with the Blue Cross hospital service associations have participated in the development of medical expense insurance, the established insurance companies have taken the lead in promoting these new forms of protection. The companies insure the majority of the persons covered, principally through group insurance written by the life insurance companies. Group insurance, as distinguished from individual insurance, is provided by means of a single master contract issued usually to an employer to insure the group of his employees and frequently their wives and children.

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

Group surgical insurance provides cash reimbursement for the amount of the surgical or obstetrical fee, subject to a maximum amount of reimbursement which is dependent on the nature of the operation. The maximum amounts of reimbursement with respect to various types of operations are set forth in a schedule that appears in the master group policy and in the certificates of coverage furnished to the individual employees. In those cases where the surgeon's charge is in excess of the maximum amount of reimbursement for the operation, the insured employee receives the maximum reimbursement from the insurance company and pays the excess from his own funds. Coverage of obstetrical procedures is optional but is usually included.

A typical schedule of maximum reimbursement will be found to provide a maximum reimbursement of $150 for the most serious operations, such as cutting into the cranial cavity, maximum reimbursements of $100 for an appendectomy, $50 for an operation for a single hernia, $25 for a tonsillectomy, and $50 as the obstetrical benefit for a natural delivery.

Although group surgical insurance is a form of group medical expense insurance, the term "group medical expense insurance" has come to be used to describe the plans of group insurance covering the costs of medical services for other than surgery and obstetrics. A wide variety of such plans with differing benefits is available. Reimbursement may be provided for doctors' visits, usually in the amount of $2 for visits by the patient to the doctor's office or by the doctor at a hospital, and in the amount of $3 for visits by the doctor to the patient's home. Coverage may be comprehensive and include reimbursement for the expense of all visits; sometimes the first one or two visits in any illness are excluded. More limited coverage plans are also being written. Under total disability plans, reimbursement is made only for the doctor's visits needed in illnesses or injuries which are disabling; under in-hospital plans, only those visits in medical cases requiring hospitalization are reimbursed. The coverages for doctors' visits may be supplemented by additional coverage for the cost of diagnostic x-ray and laboratory examinations, in accordance with a schedule of maximum reimbursement for such examinations analogous to the schedule of maximum reimbursement for surgical operations.

These new forms of insurance involve numerous insurance and statistical problems. This is equally true whether the coverage is being written by an established insurance company or by a community plan, similar to the Blue Cross, although sometimes the sponsors of such plans do not appreciate this fact as fully as they might.

The first problem is that of securing some basis for the establishment of the initial premium rates. The basic factors required for the calculation of premium rates are: first, a rate or frequency of occurrence of the event insured against, for example, the annual rate of surgical procedures per hundred insured; and, secondly, the average amount of the benefit payable when the event occurs.

While there were no insurance data directly applicable to the calculation of premium rates for group surgical insurance, experience developed in connection with hospitalization and other forms of insurance did furnish collateral information that was helpful. One company computed a scale of premium rates for group surgical insurance as follows.

At first, surgical coverage was limited to those cases where the surgery was performed in a hospital. This made it possible to base the annual frequency of surgical procedures on the annual rate of hospitalization, which experience showed to be 7 per hundred for male employees. It was observed that $\frac{3}{4}$ of all hospitalized cases involved surgery and accordingly the annual frequency of surgical procedures for male employees became 5.3 per hundred. Then, after study of the proposed schedule of benefits, it was estimated that $50 would be the average amount of each claim. The product of the frequency and the average benefit gave $2.65 for the net annual claim cost for a male employee. This figure, when increased by the amount calculated to be necessary for expenses of operation and contingencies, resulted in a gross monthly premium rate of $.35 per male employee.

The gross monthly premium rate for a female employee was set at $.70, double that for a male employee. This was consistent with the experience under other forms of disability insurance; including both weekly indemnity benefits and hospitalization insurance, where the claim cost of employed females was observed to be twice that of males. In practice, separate premium rates were not charged for male and female employees but a composite rate was computed at the effective date of the group policy depending on the percentage of female employees covered. If less than 11 per cent of the coverage was on female employees, the premium rate was not increased above the male rate of $.35. If more than 11 per cent and less than 21 per cent of the coverage was on female employees, the basic male rate was increased by 15 per cent. If the percentage of coverage on females was more than 21 per cent and less than 31 per cent, the basic male rate was increased by 25 per cent and so on.

Insurance was also to be offered for the wives and children of employees, and so a separate premium rate had to be calculated for family coverage. A gross monthly cost of $.70 was arrived at for nonmaternity surgery for a dependent wife and a cost of $.60 for obstetrical benefits. For a child, the gross monthly cost was estimated to be the same as for a male employee, namely $.35. It was decided to charge the same premium for each family covered regardless of how many children the family included; and for this purpose it was assumed that, on the average, a dependent family unit would be composed of a wife and two children. Thus, the gross monthly premium rate for dependents became $1.40 per family unit where obstetrical benefits were not to be included and $2.00 per family unit where obstetrics were to be covered.

The initial basic monthly premium rate for employee coverage was subsequently increased from $.35 to $.40 when the benefits were extended to cover out-of-hospital as well as in-hospital surgery and has since remained unchanged. The rates for family coverage were progressively reduced as experience developed and it was seen that they were redundant. The present gross monthly rates for the same company are $1.20 per family unit without obstetrical benefits and $1.55 including obstetrical benefits.

The construction of initial premium scales for the group medical expense insurance plans was similar to the calculation described in connection with surgical benefits. However, the statistical basis was even less satisfactory. It was necessary to fall back to a greater extent on studies made of the utilization of medical care by uninsured individuals. What there were of such studies were usually quite limited in their usefulness. Generally, the number of persons under observation was small the locale of the group being investigated was restricted and the amount of pertinent information secured was meager. Furthermore, since these studies were undertaken mostly from the point of view of the economic theorist or sociologist, they are not very easily adapted to the requirements of the insurance actuary.

A notable exception was a study conducted during the period from 1928 to 1931 on a sample of 9,000 families, chosen geographically and in income groups so as to be representative of the white population of the nation. Each family was interviewed at intervals over a twelve months period in order to secure detailed information regarding the incidence of illness and the extent and kind of medical care sought in connection with it. The statistical data were analyzed and published in the United States Public Health Reports and also were used extensively in the series of monographs published by the Committee on the Costs of Medical Care.

Much caution has to be exercised in using investigations based on uninsured lives for insurance purposes. The methods of collecting and classifying the data must be reviewed, because, for example, an enumeration process such as employed in the study of 9,000 families tends to underestimate the true figures. The period covered by such a study has also to be considered in the light of the rapidity with which hospital and medical practices are changing.

Of more importance in viewing studies among uninsured persons is the fact that the very presence of insurance benefits materially influences utilization. A ready example is seen in the growth of hospitalization insurance as one of the factors contributing to the overcrowding of hospitals. The use of hospital and medical facilities is, to a large extent, under the control of the individual and his physician. It is true that the degree of control is less for hospital confinement and surgical operations than for home and office calls or diagnostic x-rays. Nevertheless, an increase in utilization rates is to be anticipated under any hospital or medical expense insurance plan; this is a usual consequence of the reduction of the cost which would otherwise be a deterrent to the demand for services and their use. The greater the degree of control on the part of the insured and the doctor, the greater this effect may be expected to be. Moreover, experience demonstrates that each group insurance coverage tends to produce its own characteristic rates of utilization and related statistical factors. For one thing, the claim experience is affected by the practices of the employers and the insurance carrier in administering the insurance contracts.

After a plan of insurance of the type being discussed has been developed and has begun to be sold, the next questions are: how is the experience to be watched and what statistics are to be compiled? Since these plans do not involve the accumulation of funds to offset the future cost of deferred benefits, as in the case of life insurance or annuities, a simple and very effective index of experience may be obtained by a comparison of premiums which become due with the claims that are incurred. Such a comparison brings out very readily the financial effect of the plan on the company's operations and is easy to compute from month to month and to maintain on a current basis. However, such ratios do little to satisfy the need for more detailed statistics for rate making and other similar purposes.

Group insurance policies are written for a period of one year and are renewable for annual periods thereafter. The premium rates at issue are guaranteed for the first policy year but are subject to re-rating upward or downward by the insurance company at the end of that period and at the end of each succeeding policy year. The policies participate in the surplus earnings of the company annually through the declaration of dividends or the allowance of retroactive rate credits, where earned.

Involved in the computation of dividends or retroactive rate credits is the process known as experience-rating, whereby weight is given both to the claim experience of the individual group policy and to the claim experience of the class of policies to which it belongs. It is clear that in order to make possible renewal underwriting and the computation of dividends or retroactive rate credits, a record of premiums due and the number and amount of claims incurred must be kept separately for each group policy for each policy year.

It is natural then that these individual group experience records should be the basis of further statistical analysis within a company. Each group policy can be classified in a number of ways: separately as to the coverage of employees and family units; according to the basic type of coverage it provides; according to the variations in specific benefit provisions, such as whether the maximum period of coverage for hospital expenses is 31 days or 70 days in any one confinement; according to the percentage of female employees covered; according to the industry of the employer, and by geographical location. Thus, a compilation of data regarding claims and premiums by policy years makes possible an extensive statistical analysis. Annual dollar claim costs per unit of benefit can be produced, as well as annual claim frequencies per insured person or family, and the average amounts of benefits paid per claim. These can be obtained separately for male employees, for female employees and for family units. In addition, the variation in these factors by industry, policy duration, calendar year and geographical location may be studied.

This further analysis still does not give all the statistical information necessary for various purposes because it does not involve the classification of individual claims other than by the type of coverage and the group policy under which they were incurred. No information is furnished concerning the variation in claim frequencies, for example, by age, by cause of disability or type of surgical operation, or by duration of disability or period of hospitalization. Only the appropriate classification of individual claims upon their completion and the analysis of their relationship to the numbers exposed to risk can produce such information accurately.

Here arises a question of approach about which there may be some difference of opinion among the insuring companies. The question is whether there should be set up, on a continuing basis for statistical purposes, a system of coding individual claims, involving the use of all or a specific fraction of all claims. The alternative is, of course, to develop the basic claim rates and claim costs on a continuing basis but to rely on occasional sample studies designed for specific purposes for any additional detailed information which might be desired.

The merits of the question revolve in good part around the cost of the two approaches and can be brought into focus by mentioning that currently one company is paying annually about 75,000 hospital expense claims, 65,000 surgical claims and 45,000 medical expense claims. In the writer's opinion, the additional detailed information does not warrant the cost of securing it on a continuing basis as against the lesser cost of conducting periodic or occasional sample studies. Of course, some sample studies would have to be made anyway because the wide variety of information desired from time to time would not all justify inclusion in any continuous system.

So far, the discussion has been in terms of the operations of a single company. Whenever possible, it is desirable to improve the reliability of statistical results by increasing the body of data available for study. Thus, the combined experience of a number of companies is not only more reliable than that of a single company but it also furnishes a standard by means of which any one company may judge the efficiency of its operations. The medium for the compilation and analysis of group insurance experience on an intercompany basis is a standing committee of the Actuarial Society of America known as the Committee to Prepare Mortality and Morbidity Studies on Group Insurance.

The work of this Committee was interrupted by the war, as were many other statistical efforts of the individual companies. Accordingly, it is only now, after discussions held during the past 18 months, that the Committee has completed plans for the compilation of intercompany group morbidity experience. The contributing companies will be furnishing the Committee with aggregate "claims" and "exposed to risk" for each group policy appropriately classified so as to enable the derivation of unit claim costs. It is expected that the Committee's first report on group morbidity experience will be published in the 1948 *Transactions* of the Actuarial Society. For the time being, group medical expense insurance is not being studied, largely because of the newness of the coverage, the wide variety of plans being written, and the diversity in policy provisions. It is hoped that it will be possible to commence studying group medical expense insurance on an intercompany basis at a later date.

The Committee has also undertaken a sample study based on individual surgical claims. The study is designed to supply information regarding, first, the relative frequency with which different surgical procedures occur and, secondly, the relative values of the fees being charged by doctors for performing such procedures. There is little known about either of these matters. A detailed frequency distribution of various types of operations should be on hand for the computation of premiums for a given surgical schedule in its entirety, as well as to enable the computation of the cost of a variation in benefit that might be under consideration for a particular surgical operation.

With respect to doctors' actual charges, the companies are interested not so much in their dollar value since it is not contemplated that the surgeon's fee will necessarily be covered in full by the insurance. The companies are concerned instead, with the relationships which doctors' charges for different operations bear to one another. For, if surgical insurance is to be most effective, it is important that the maximum amount of reimbursement provided for each type of operation be reasonably proportionate to the surgeon's fee. Thus, it is not desirable to have a surgical schedule which might on the average insure the full cost of tonsillectomies, for example, while covering only a small fraction of the average surgeon's charge for performing an appendectomy. A scale of the relative values of actual charges will make it possible to determine whether existing surgical schedules are in accord with the pattern of the fees being charged by physicians, and will permit the revision of the schedules where they may not be.

In planning the intercompany study, one of the problems encountered was that of devising a suitable classification code for the various surgical procedures. A code was desired that would be neither too detailed to be practical nor insufficiently detailed for adequate analysis. The information was to be coded from claim papers which do not always give complete medical descriptions. Furthermore, the bulk of the coding work would be done by clerks whose medical knowledge was limited, although the assistance of the insurance company doctors would be available. With the advice of several company doctors, a satisfactory code containing about 400 classifications was prepared.

In order that the study of the relative frequencies of various surgical procedures be free of seasonal bias, it was necessary that the claims included in that portion of the study be chosen proportionately over a twelve months' period. Several of the companies found it inconvenient to select claims in that way and so are participating only in the study of the relative charges made by doctors.

In all, punched cards covering about 100,000 claims have been prepared by seven companies. The following information has been submitted with respect to each surgical claim: (1) the code classification of the first surgical operation performed, (2) the classification of a second operation which may have been performed at the same time, (3) the total number of procedures performed during the same operation, (4) whether the operation was performed in a hospital or elsewhere, (5) the claimant's sex, including whether an employee or dependent, age and state of residence, (6) the total amount of the doctor's charge, and (7) the amount of the insurance benefit payable. The data from the seven companies have been received by the compiling company and are in the process of analysis. The indications are that the study[1] will produce much information useful not only to insurance companies but also to others.

In concluding, it should be emphasized that insurance against the cost of medical care is relatively new. Much is yet to be learned, especially about plans insuring doctors' charges for non-surgical services; these are very definitely in the formative, experimental stages. The insurance companies are finding out what demand there is for various medical services, what it costs to provide such services and what administrative problems are involved in covering those costs on a prepayment basis. It is hoped that the statistics and experience of the insurance companies will contribute valuable knowledge regarding the important and difficult problem of distributing and adequately paying for medical services in this country.

---

[1] See "Group Surgical Claims Study, Report of the Committee to Prepare Mortality and Morbidity Studies on Group Insurance," *Transactions* of the Actuarial Society of America, Vol. 49, 1948.

BIBLIOGRAPHY

Collins, S. D., Series of papers based on records for 9,000 families, *United States Public Health Reports*, various, volumes 48–57.

Falk, I. S., M. C. Klem and N. Sinai, *The Incidence of Illness and the Receipt and Costs of Medical Care among Representative Families, 1928–1931*, Pub. No. 26, Committee on the Costs of Medical Care, University of Chicago Press, 1933.

Fitzhugh, G. W., "Further Developments in Group Hospital Expense Insurance," *Record* of the American Institute of Actuaries, vol. 29, 1940, p. 224.

Hunter, A. and A. B. Thompson, "Hospital Service Insurance," *Transactions* of the Actuarial Society of America, vol. 44, 1943, p. 5.

Milliman, W. A., "Insurance of the Expense of Medical Service," *Transactions* of the Actuarial Society of America, Vol. 41, 1940, p. 114.

Milliman, W. A., "Insurance Carriers and Medical Care Plans," *Health Insurance in America;* Address at the Second National Conference on Social Security, January 1945.

Reed, L. S., *Blue Cross and Medical Service Plans*, U. S. Public Health Service, Federal Security Agency, Washington, D. C., 1947.

van Steenwyk, E. A., "The Administration and Underwriting of Hospital and Medical Insurance," *Health Insurance in America;* Address at the Second National Conference on Social Security, January 1945.

PROCEEDINGS

107th ANNUAL MEETING

HOTEL COMMODORE, NEW YORK CITY

PROGRAM

Sunday, December 28, 1947

—9:00 A.M.—

REGISTRATION

—10:00 A.M.—

STATUS AND PROSPECTS IN METHODS OF SOCIAL RESEARCH

(Joint with the American Sociological Society)

Chairman: Raymond V. Bowers, Research and Development Board

Discussion: Stuart C. Dodd, University of Washington
Philip M. Hauser, University of Chicago
Talcott Parsons, Harvard University

THE TEACHING OF STATISTICS

(Joint with the Institute of Mathematical Statistics)

Chairman: William G. Cochran, Institute of Statistics, University of North Carolina

A Review of the Present Situation and Recommendations
A Digest of Recent Reports
James D. Paris, Metropolitan Life Insurance Co.
(Prepared for the American Statistical Association Temporary Committee on Training of Statisticians)

Discussion: Robert W. Burgess, Western Electric Co.
Burton H. Camp, Wesleyan University
Joseph F. Daly, Bureau of the Census
Philip J. Rulon, Harvard Graduate School of Education

PROGRESS IN THE INDUSTRIAL CLASSIFICATION OF ESTABLISHMENTS

Chairman: Emil D. Schell, Bureau of Labor Statistics

Introductory Remarks, V. S. Kolesnikoff, Bureau of the Budget
The Revised Standard Industrial Classification Structure
Frank Cizar, Bureau of the Budget
The Applications of the Standard Structure
Margaret Martin, Bureau of the Budget

The Coordination of Federal and State Social Security Classifications
William Cummins, Bureau of Old Age and Survivors Insurance
Frank Nagel, Bureau of Employment Security

CLASSIFICATION PROBLEMS IN THE 1947 CENSUS OF MANUFACTURES

Discussion: Meredith Givens, New York State Department of Labor
Samuel P. Hayes, Jr., Dun and Bradstreet
Loring Wood, Bureau of the Budget

—1:15 P.M.—

STATUS AND PROSPECTS IN POPULATION PLANS

(Joint with the American Sociological Society)

Chairman: Philip M. Hauser, University of Chicago

Discussion: Frank Notestein, Princeton University
Henry S. Shryock, Jr., Bureau of the Census
Irene Taeuber, Princeton University
P. K. Whelpton, Scripps Foundation for Research in Population Problems

ESTABLISHMENT SAMPLING—A ROUND TABLE

Chairman: Lester S. Kellogg, Bureau of Labor Statistics
Lester Frankel, Dun and Bradstreet
Margaret Gurney, Bureau of the Census
Samuel Weiss, Bureau of Labor Statistics

THE TEACHING OF STATISTICS

(Joint with the Institute of Mathematical Statistics)

Chairman: Francis G. Cornell, University of Illinois

Syllabus for a Proposed Course in Basic Statistics
George W. Snedecor, Iowa State College

Discussion: Elmer B. Mode, Boston University
Samuel A. Stouffer, Harvard University
Helen M. Walker, Teachers College, Columbia University
Albert E. Waugh, University of Connecticut
Film: The Presentation of a film on Modern Quality Control
Simon Collier, Director of Quality Control, Johns Manville Company

—3:30 P.M.—

INTERNATIONAL PROBLEMS IN HEALTH AND VITAL STATISTICS

(Biometrics Section of the American Statistical Association Joint with the Biometric Society)

Chairman: Halbert L. Dunn, National Office of Vital Statistics

Principles Underlying the Sixth Decennial Revision of the International Lists of Diseases and Causes of Death and their Potential Uses
Lowell J. Reed, Johns Hopkins University

Review of Work and Actions of the Expert Committee for the Preparation of the Sixth Decennial Revision of the International Lists of Diseases and Causes of Death of the World Health Organization
W. Thurber Fales, Baltimore City Health Department
Tabulations Needed for International Purposes in the Fields of Health and Vital Statistics
Forrest E. Linder, Statistical Office of the United Nations
The Census Base for Satisfying Health and Vital Statistics Needs
Calvert L. Dedrick, Bureau of the Census

PATTERNS OF WORKING LIFE

Chairman: Gladys Palmer, University of Pennsylvania

Labor Force Life Tables
Seymour L. Wolfbein, Bureau of Labor Statistics
Experience Patterns in Covered Employment
J. Pearlman, Bureau of Old Age and Survivors Insurance
Patterns of Occupational Mobility
Lloyd Reynolds, Yale University

Discussion: Margaret Hagood, Bureau of Agricultural Economics
A. J. Jaffe, Bureau of the Census
Paul Webbink, Social Science Research Council
Emmett H. Welch, Bureau of the Census

MEASUREMENT AND SIGNIFICANCE OF INVENTORIES

Chairman: Ralph J. Watkins, Dun and Bradstreet

Measuring Physical Inventories
Weir M. Brown, Board of Governors of the Federal Reserve System
Significance of Inventories in the Economic Picture
Louis J. Paradiso, Econometric Institute

Monday, December 29, 1947

—9:00 A.M.—

OPERATIONAL RESEARCH

(Joint with the Institute of Mathematical Statistics, with the cooperation of the Operations Evaluation Group of the Navy and Operations Analysis of the Air Force)

Chairman: Edward L. Bowles, Massachusetts Institute of Technology

Operations Research in the Department of the Navy
J. Steinhardt, Director, Operations Evaluation Group
Operational Research in the Department of the Air Force
Leroy A. Brothers, Chief, Operations Analysis

Discussion: Arthur A. Brown, Operations Evaluation Group
Thomas I. Edwards, Operations Analysis
Galen B. Price, University of Kansas
W. J. Youden, Douglas Aircraft Company

MORE ADEQUATE CURRENT REPORTS ON INDUSTRIAL ACTIVITY

Chairman: David Novick, University of Puerto Rico

Speakers: Maxwell Conklin, Bureau of the Census
Edward Hincks, Bureau of the Budget
Geoffrey H. Moore, National Bureau of Economic Research
Nathan Weinberg, United Automobile Workers, Congress of Industrial Organizations

Discussion: Donald Kershner, Crown Central Petroleum
Victor Perlo, Monetary Research Division, U.S. Treasury Department
Lorman Trueblood, Board of Governors, Federal Reserve System

—9:30 A.M.—

SOME ASPECTS OF GRAPHIC PRESENTATION

Chairman: Kenneth W. Haemer, American Telephone and Telegraph Co.

Elements of Presentation
H. C. Barton, Jr., Presentation, Inc.
The Place of Graphics in the Teaching of Statistics
Henry S. Miller, Queens College
Production of Display Presentation
Charles J. Hauk, Jr., Chartmakers, Inc.

—10:00 A.M.—

STATISTICAL PROBLEMS IN HIGHWAY PLANNING AND TRANSPORT RESEARCH

Statistics in Highway Planning and Transport Research
John T. Lynch, Public Roads Administration
Factors Affecting Generation and Distribution of Highway Traffic
Nathan Cherniack, Port of New York Authority
Application of Poisson Distribution Theory to Traffic Control
Donald Schapiro, Yale Bureau of Highway Traffic
Analysis of Motor Vehicle Registration Records
G. P. St. Clair, Division of Financial and Administrative Research, Public Roads Administration

—11:00 A.M.—

THE CURRENT STATUS OF FEDERAL STATISTICS—A ROUND TABLE

Chairman: Stuart A. Rice, Division of Statistical Standards
Ewan Clague, Bureau of Labor Statistics
Howard C. Grieves, Bureau of the Census
Orris V. Wells, Bureau of Agricultural Economics

—12:00 M.—

COMMITTEE ON CENSUS ENUMERATION AREAS—*Luncheon*

Chairman: Howard Whipple Green, Cleveland Health Council

Plans for the 1950 Population Census Related to Census Tracts and Block Data
Leon E. Truesdell, Bureau of the Census

Tracted Cities in Canada
O. A. Lemieux, Dominion Bureau of Statistics, Ottawa, Canada
Use of Data by Blocks in St. Louis
Edward B. Olds, Metropolitan St. Louis Census Committee
General Discussion: Series of Four Minute Reports from Census Tract Key Persons

—1:15 P.M.—

SOCIAL PSYCHOLOGY

(Joint with the American Sociological Society)

Chairman: Herbert Goldhamer, University of Chicago

Discussion: Joseph K. Folsom, Vassar College
Abraham Kardiner, New York City
Richard T. LaPierre, Stanford University
C. Wright Mills, Bureau of Applied Social Research

MORE ACCURATE INDUSTRIAL STATISTICS

Chairman: David Novick, University of Puerto Rico

Melvin Anshen, Indiana University
Solomon Barkin, Textile Workers Union
Ewan Clague, Bureau of Labor Statistics
F. W. Heintzelman, American Institute of Accountants

Discussion: G. A. Renard, National Association of Purchasing Agents
Boris Shishkin, American Federation of Labor
George A. Steiner, University of Illinois

—3:15 P.M.—

BIOMETRICS SECTION—CONTRIBUTED PAPERS

Chairman: J. H. Watkins, Yale University

Triple Rectangular Lattices
Boyd Harshbarger, Virginia Polytechnic Institute
Estimation of the Mean and its Error from Grouped Poisson Distributions
C. I. Bliss and M. C. Jackman, Connecticut Agricultural Experiment Station
On the Use of Parallel or Nonparallel Systems of Transformed Curves in Bio-assay
William R. Thompson, New York State Department of Health
Observations on Cost Utility Rating in Relation to Correlation
Joseph Berkson, Mayo Clinic
"Use of Statistical Methods in Human Genetics." *A Test for Monomeric Inheritance of a Normal Trait in Man*
Marianne E. Bernstein, Syracuse University

—3:30 P.M.—

OPERATIONAL RESEARCH AND OTHER MILITARY APPLICATIONS OF STATISTICS

(Joint with the Institute of Mathematical Statistics, with the cooperation of the Operations Evaluation Group of the Navy and Operations Analysis of the Air Force)

Chairman: Merrill M. Flood, General Staff, United States Army

Operations Analysis in the Southwest Pacific Air War
Roger I. Wilkinson, Bell Telephone Laboratories
Operations Analysis of Air-Sea Rescue
E. S. Lamar, Operations Evaluation Group
Factorial Chi-Square in Test Shooting
A. E. Brandt, Naval Ordnance Laboratory
Mathematical Techniques of Program Planning
George Dantzig, Consultant to the Air Comptroller, Headquarters, United States Air Force

MEASURES OF CONSTRUCTION ACTIVITY

Chairman: Paul Krueger, Bureau of the Budget

A User's View
Leo Grebler, Private Consultant
The Government's Estimates of Residential Building
H. E. Riley, Bureau of Labor Statistics

Discussion: Howard Whipple Green, Cleveland Health Council
W. A. Berridge, Metropolitan Life Insurance Co.
Samuel Dennis, Department of Commerce
W. G. Shaw, Department of Commerce
Clyde Shute, F. W. Dodge, Inc.

INSTITUTE OF MATHEMATICAL STATISTICS—BUSINESS MEETING

—5:30 P.M.—

AMERICAN STATISTICAL ASSOCIATION—ANNUAL BUSIENSS MEETING

—8:00 P.M.—

PRESIDENTIAL ADDRESSES

(Joint with the American Sociological Society)

Willard L. Thorp, President, American Statistical Association
Louis Wirth, President, Americal Sociological Society

Tuesday, December 30, 1947

—9:00 A.M.—

HOUSING

(Joint with the American Sociological Society)

Chairman: Howard Brunsman, Bureau of the Census

An Appraisal Method for Measuring the Quality of Housing
Allan A. Twichell, American Public Health Association Committee on the Hygiene of Housing
Problems in Development and Application of Standards of Housing in Making Adequate Budgets
Dorothy S. Brady, Bureau of Labor Statistics

Discussion: John P. Dean, Research Fellow, Social Science Research Council
Nicholas J. Demerath, University of North Carolina

—10:00 A.M.—

DESCRIPTIVE STATISTIC IN THE SERVICE OF LIFE INSURANCE

Chairman: Joseph B. MacLean formerly Vice President and Actuary of the Mutual Life Insurance Company of New York and past President of the Actuarial Society of America

Life Insurance Investments and the Capital Markets
Fergus J. McDiarmid, Lincoln National Life Insurance Company
Mortality Statistics for Life Insurance Underwriting
Edward A. Lew, Metropolitan Life Insurance Company
Statistical Problems in New Fields of Insurance
Morton D. Miller, Equitable Life Assurance Society

—9:30 A.M.—

APPLICATION OF THE THEORY OF EXTREME VALUES
(Joint with the Institute of Mathematical Statistics)

Applications to Fracture Problems
Benjamin Epstein, Carnegie Institute of Technology
Applications to the Prediction of Flood Flows
Emil Gumbel, Brooklyn College
The Mathematical Theory of Extreme Values
Richard von Mises, Harvard University
Applications to Meteorology
Horace Norton, Weather Bureau

—11:00 A.M.—

PROGRESS REPORT AND DISCUSSION OF PRODUCTIVITY IN INDUSTRY
Ernest Dale, Columbia University, American Management Association
Katherine P. Ellickson, Congress of Industrial Organizations
W. D. Evans, Bureau of Labor Statistics
Martin Gainsbrugh, The Conference Board
Thomas Mills, Bureau of the Budget
Lazare Teper, International Ladies Garment Workers Union

—1:00 P.M.—

OUTLOOK FOR 1948

Chairman: Willard L. Thorp, President, American Statistical Association

Discussion: Martin Gainsbrugh, The Conference Board
James Hughes, Auchincloss, Parker and Redpath
Ragnar D. Naess, Naess and Commings
Imrie de Vegh, Consulting Economist

BIOMETRICS SECTION—BUSINESS MEETING

STATISTICAL TECHNIQUES IN LIFE INSURANCE

(Joint with the Institute of Mathematical Statistics)

Chairman: Robert J. Myers, Social Security Administration

Problems with Sampling Procedures for Reserve Valuations
George C. Campbell, Metropolitan Life Insurance Co.
Sampling Errors in Life Insurance Mortality and Other Statistics
Donald Cody, Equitable Life Assurance Society
Recent Developments in Graduation and Interpolation
T. N. E. Greville, National Office of Vital Statistics

—3:30 P.M.—

CENSUS AND VITAL STATISTICS

(Joint with the American Sociological Society)

Chairman: Edmund de S. Brunner, Columbia University

Sociologists' Suggestions Relating to the 1950 Census in the U.S.
P. K. Whelpton, Scripps Foundation for Research in Population Problems
Comments on Feasibility of the Sociologists' Suggestions
A. Ross Eckler, Bureau of the Census
Problems in Revision of the Birth and Death Certificates
Felix Moore, National Office of Vital Statistics
Discussion: T. Lynn Smith, Vanderbilt University
Dorothy Swaine Thomas, Giannini Foundation

THE USE OF STATISTICS BY BUSINESS MANAGEMENT—A ROUND TABLE

Chairman: Isador Lubin, United Nations Labor and Employment Commission
Robert W. Burgess, Western Electric Co.
W. H. Garbade, Shell Oil Co.
Richard O. Lang, S. C. Johnson & Son, Inc.
Charles Roos, Econometric Institute
Clarence Smith, International Business Machines

AMERICAN STATISTICAL ASSOCIATION—CONTRIBUTED PAPERS

Factors in the Discontinuance of a Statistical Series
Harry Alpert, Bureau of the Budget
Testing for Grade and School Differences Among the Scores of Home Economics Students
William D. Baten, Brooklyn College
On Measuring Languages for International Use: On Pan Sampling
Stuart C. Dodd, University of Washington Public Opinion Laboratory
Sampling Errors and Components of Interviewer Costs in Survey Design
Roe Goodman and Charles Cannell, Survey Research Center
The Sugar Consumption Quota: A Problem in Administration of Legislative Controls
Julius Guttenplan, Sucrest Corp.
Income, Savings, and Mortgage Investment
Mortimer Kaplan, Federal Housing Administration

Objective Selection of a Respondent within the Household
Leslie Kish, Survey Research Center, University of Michigan
Trends in Statistical Activities of Trade Associations
Sidney Lieberman, Chart Facts
Maximum Likelihood Estimates of Absolute Dispersion for Samples Drawn from Type III and Type V Populations
Nilan Norris, Hunter College

INSTITUTE OF MATHEMATICAL STATISTICS—CONTRIBUTED PAPERS

THE BIOMETRIC SOCIETY, EASTERN NORTH AMERICAN REGION—BUSINESS MEETING

List of Committee Chairmen and Representatives for 1947

Committee on Fellows: Joseph S. Davis
Committee on Nominations: Woodlief Thomas
Committee on Graphics: Kenneth W. Haemer
1947 Annual Meeting Program Committee: Samuel S. Wilks
International Statistical Conferences ASA Program Committee: Willard L. Thorp, Philip M. Hauser, *Acting Chairman*
Committee on Census Enumeration Areas: Howard W. Green
Joint Committee on Occupational Classification: Gladys L. Palmer
Census Advisory Committee: W. F. Ogburn
Advisory Committee to BLS on Employment and Statistics Program: S. S. Wilks
Advisory Committee to the National Housing Agency on Housing Statistics: Howard G. Brunsman
Biometrics Section Committee: D. B. DeLury
Temporary Committee on the Training of Statisticians: F. G. Cornell
Representatives on the Committee for the Preparation of a Source Book of Economic Statistics: Walter Mitchell, Jr., J. Frederic Dewhurst
Representative on the Board of Directors of the National Bureau of Economic Research: Frederick C. Mills
Members of the Social Science Research Council: Joseph S. Davis, Harold Hotelling, Willard L. Thorp
Representative on the Council of the American Association for the Advancement of Science: Frederick F. Stephan
Representatives on the Joint Committee for the Development of Statistical Applications in Engineering and Manufacturing: Churchill Eisenhart, Walter A. Shewhart
Representative on the Sectional Committee on Standards for Graphic Presentation of the American Standards Association Sponsored by the American Society of Mechanical Engineers: Kenneth W. Haemer
Representative on the Advisory Board of the American Year Book: Frank W. Notestein
Representatives on the Intersociety Committee on Science Foundation Legislation: Lester S. Kellogg, Frederick F. Stephan
Member of the American Documentation Institute: Lester S. Kellogg

MINUTES OF THE ANNUAL BUSINESS MEETING

The American Statistical Association convened for its 107th Annual Business Meeting at the Hotel Commodore, New York City, Main Ballroom, on Monday, December 29, 1947, President Thorp presiding.

The minutes of the 106th Annual Business Meeting, held in Atlantic City on January 25, 1947 were approved.

Aryness Joy Wickens, acting for Joseph Davis, Chairman of the Committee on Fellows, read the following citations of the Fellows elected for 1947:

> The Committee on Fellows has unanimously elected six members as Fellows of the American Statistical Association:
>
> *John Hamilton Curtiss*, Assistant to the Director of the National Bureau of Standards; contributor to statistical theory and to pure mathematics; teacher, administrator, and patriotic innovator in the application of statistical methods to problems of the Navy at war.
>
> *Milton Gilbert*, Chief of the National Income Division of the Bureau of Foreign and Domestic Commerce of the U. S. Department of Commerce, and formerly Editor of the *Survey of Current Business*. He has been responsible for much of the improvement of statistics of gross national product and has contributed notably to the analysis of the national income.
>
> *Morris H. Hansen and William Hurwitz*. U. S. Bureau of the Census, a distinguished statistical team which, although among the younger group of outstanding members of this society, has made important contributions to the use of sampling techniques in the recurrent report of the Census Bureau on population and unemployment. They are joint developers of large-scale techniques whereby gamblers' tools can serve solid citizens through efficient gathering of pertinent information.
>
> *Jacob Marschak*, Research Director of the Cowles Commission and Professor of Economics at the University of Chicago; born in Russia, with academic degrees from Heidelberg and Oxford, he has made valuable personal contributions and exercised constructive leadership in statistics, economics, and econometrics in England and in the United States.
>
> *Herbert Marshall*, Dominion Statistician of Canada, Chairman of United Nations Statistical Commission, noted for his responsible leadership in and important contributions to Canadian statistics, especially those on the Canadian international balance of payments and on Canadian-American direct investments; and for his services in the use of statistics in the control of his country's economy during a great war.

Lester S. Kellogg, Secretary of the Association, reported briefly on the current status of finances and membership of the Association.

The following resolution, submitted by A. Ford Hinrichs, was read:

> RESOLVED that:
>
> 1. The President of the American Statistical Association shall arrange in consultation with the President of the American Economic Association that the annual meetings in 1949 and later years shall be held at a common time and place.
> 2. In order to contribute to a unified attack on common problems, other

societies of the social sciences should be invited to join in the arrangements for a common meeting place.

3. The President of the American Statistical Association shall appoint a Committee not later than March, 1948, to arrange a program on the PROBLEMS OF THE STATISTICAL PROFESSION and of a unified attack on problems in the field of the social sciences, to constitute the core of the sessions of the annual meetings in 1949. The Committee shall have discretion as to the contents and arrangement of this program, but it is desirable that it should include a discussion of:

a) professional standards and the professional requirements of various fields of work and types of occupations in which statisticians are commonly engaged.
b) training requirements and methods, with reference not only to professional statisticians but also with reference to the elementary levels for citizenship and active participation in a democracy.
c) kind of professional organization needed to encourage the development within the total profession of groups of specialized professional interests.
d) methods of presenting the findings of the profession to its members and of presenting these findings and representing the profession to the public.
e) proposals for the development of a better coordinated attack by the profession on unsolved problems of statistical theory, including
    i) the development of a moving 5-year program to establish the goals of the Association itself,
    ii) a continuing analysis in each field of theory of the gaps in the data or in professional knowledge,
    iii) arrangements to encourage research and analysis particularly at these critical points,
    iv) arrangements at the annual meetings to present a comprehensive brief view of all new advances to the profession, with special reference to advances in statistical theory, in the measurement of relationships between variables, in data and in techniques of analysis.

While this program shall constitute the core of the annual meeting in 1949, it is not intended to exclude other types of discussion.

4. The Committee so appointed shall invite the participation of the Social Science Research Council in its deliberations and program arrangements and shall consult with committees that may be appointed for similar purposes by other associations of the social sciences. As far as possible, joint programs shall be arranged with such other associations to deal with common or similar problems.

This resolution was referred to the Board of Directors for its consideration. (A similar resolution was presented at the annual meeting of the American Economic Association and the American Sociological Society.)

Willard L. Thorp, President reporting for the Committee on Structure, stated that the draft of the new Constitution was in process of revision and would be submitted to the members within a short time. He informed the membership that through an oversight of the Board of Directors a proposed amendment to the Constitution, approved by the membership in 1940, was never resubmitted to the membership because of the failure to hold annual meetings during the recent world war. This report was noted as a matter of information and no additional vote was taken.

In the absence of Francis G. Cornell, Chairman of the Temporary Committee on the Training of Statisticians, Helen M. Walker, a member of the Committee, reported that the Committee had organized several sessions at the Annual Meeting which had been very well attended. Over 70 people present at these sessions on training signed a petition for the formation of a Section on the Training of Statisticians, and Professor Walker announced that an organization meeting to form the Section would be held the following day.

Samuel S. Wilks reported that a plan had been worked out for the publication of a new periodical to be sponsored by the Association, to be called "Statistical Reviews." The publication will begin its appearance in 1949 and will carry book reviews, abstracts, and the Statistical Methodology Index which at present appears in the *Journal*. He reported that plans had been completed and were ready to submit to a foundation with a request for subsidy.

Dr. Wilks also reported on the activities of the Committee to advise the Bureau of Labor Statistics on its employment statistics program. This Committee prepared a detailed report which included a number of useful recommendations and was submitted to the Commissioner of Labor Statistics in August of 1947. Aryness Joy Wickens, of the Bureau of Labor Statistics, expressed the thanks of the Bureau for the service rendered by the Committee.

Willard L. Thorp presented the following slate of officers prepared by the Nominating Committee:

George W. Snedecor, President;

Frederick F. Stephan and Willard L. Thorp, Directors for the years 1948, 1949 and 1950;

A. Ford Hinrichs, Rensis Likert, Abraham Wald, and Harry R. Wellman, Vice-Presidents;

Lester S. Kellogg, Secretary-Treasurer.

The vote in favor of the slate was unanimous and the Secretary was instructed to cast one ballot for the entire list.

Willard L. Thorp, outgoing President, turned the Chairmanship of the meeting over to the newly elected President, George W. Snedecor. Professor Snedecor expressed his pleasure at the honor conferred on him and his appreciation of the diligent and unremitting efforts of Lester S. Kellogg as Secretary of the Association.

The following resolution was presented. It is RESOLVED that the Officers and the members of the Association express their appreciation to the New York Chapter for its help in making such excellent arrangements for the Annual Meeting. The thanks of the membership are particularly due to Waite S. Brush, District Representative of the New York Chapter and Chairman of the local Committee on Arrangements, for his effective assistance.

It was also RESOLVED to commend Waite S. Brush for the remarkable expansion of activities in the New York area during the past year.

Solomon Barkin presented a resolution stating that "The American Statistical Association should investigate its responsibilities in the field of industrial statistics." This resolution was referred to the Board of Directors for action.

Waite S. Brush reported that there had been a tremendous increase in activities in the New York area. He reported that the Social Statistics Section and the Statistical Techniques Section of the Chapter had been meeting regularly, that a number of full Chapter Meetings had been organized and well-attended during the course of the past year. The Chapter has been occupied with the statistical problems of the New York area and has been working with municipal and state committees to determine statistical requirements in New York and in helping to plan means of making additional statistical data available. Mr. Brush appealed for help from the national Association in obtaining the services of a paid secretary and asked the attention of the President and the Board of Directors for his request.

The meeting was adjourned.

LESTER S. KELLOGG, *Secretary*

Report of the Board of Directors

The year 1947 witnessed a continuing growth of the American Statistical Association and an expansion and ramification of its activities. The dues-paying membership increased from approximately 3,800 to about 4,800. Three new chapters were organized, in Honolulu, Ithaca and Oklahoma City. Negotiations were begun for the chartering of additional chapters, two of which were formally organized during the first quarter of 1948.

During September 1947, many of the world's leading statisticians met in Washington for the sessions of the International Statistical Conferences. Five hundred scientific workers in statistics and allied fields attended and all major countries, with the exception of the U.S.S.R., sent delegates. The groundwork was laid for the resumption of that fruitful international cooperation among scientists which characterized the pre-World War I and inter-war periods.

The Association was one of the sponsoring organizations of this gathering. In addition to giving a reception at which the foreign delegates were able to renew social and intellectual contacts which had been broken during the War, the ASA organized three of the sessions of the Conferences. A special issue of the *Journal* was prepared for the international meetings and distributed gratis to the foreign delegates. A non-commercial graphics exhibit was organized to familiarize non-United States delegates with the presentation methods developed in this country over the past seven years. Through the cooperation of publishing firms and statistical machine manufacturers, a commercial display was provided which gave comprehensive coverage of textbooks, publications on methodology and specialized fields and mechanical aids to statistical work, ranging from simple computing machines to highly advanced and specialized types. The exhibit produced a net income of $3,286.20, which was made available to the Joint Arrangements Committee to help in defraying expense incident to conducting the Conferences.

Recognizing that physical destruction of foreign libraries, dollar exchange shortages and other factors have conspired to deprive foreign statisticians and statistical institutions of the tools necessary to their work, the Association made available complete sets of the ASA *Journal* for the past 15 years gratis to approximately 25 war-devastated libraries. Current books on statistical developments and methodology were also made available to needy foreign statisticians, scholars and schools.

The International Statistical Conferences in Washington resulted in the organization of several new specialized groups of statisticians and persons working in closely related fields. International groups were launched to study national income and wealth, public opinion sampling, biometrics and demography. In all these new ventures, members of the Association have played an active role.

The papers delivered at the Conferences are to be published in five volumes. The Association is editing the contributions made to the three sessions which it organized. Noteworthy papers of special interest to the membership will be published separately in forthcoming issues of the *Journal.*

In August 1947, the first issue of *The American Statistician* appeared.This bi-monthly periodical, which supplants the organizational *Bulletin* of the Association, was planned as a professional news magazine. It carries brief reports of important statistical news, discussions of professional problems, and articles on the activities of statisticians. The magazine is designed to reach not only the members of the Association but that large group of students and users of statistics who are interested in the activities of the profession. Its circulation at the close of 1947 was around 5,400.

The 107th Meeting of the Association was held in New York City at the Hotel Commodore in December 1947 in conjunction with the American Sociological Society, the Biometric Society and the Institute of Mathematical Statistics. Despite the New York blizzard and the meetings of the American Economic Association and of Section K of the AAAS simultaneously in Chicago, the New York meetings drew an attendance of over 1,000 persons. A feature of the New York meeting was the comprehensive display of statistical presentation methods prepared for the Association by its Committee on Graphics under the chairmanship of Kenneth W. Haemer.

Further discussion of the new Constitution of the Association occurred at the December meeting in New York. The document discussed at that meeting was sent back for redrafting. This redraft was accepted, by mail ballot of the membership, in early 1948.

Plans for a joint Annual Meeting between the American Economic Association and the ASA in Cleveland in the last week of December 1948 have been endorsed by both organizations. Tentative arrangements for a joint annual meeting on the West Coast in 1949 or 1950 are currently under consideration.

The central office of the Association in Washington acquired additional space and staff in 1947 to meet the requirements of continuing ASA growth. At the close of 1947, Lester S. Kellogg, secretary-treasurer of the Association during the period of rapid expansion of the war and post-war years, accepted an executive position with Deere and Company in Moline, Illinois. In view of his absence from Washington, Sylvia C. Weyl was named Acting Secretary-Treasurer on an ad interim basis and a committee was appointed to study the question of finding a successor to Dr. Kellogg.

The development of the statistical field has required increasing activity and responsibility on the part of the Association. There is a demand for more publications, more meetings with other professional groups, more professional representation, an employment service and more committee activity. These activities require more resources and an expanded secretarial staff. Costs of printing and other services have about doubled since the end of the war. The effects of inflation have been partly offset by the growth in membership and subscriptions and the contributed secretarial and editorial services. The old membership rate of $5.00 a year was found to be inadequate to cover basic expenses much less to permit expanded activity. In view of this situation the membership was asked to increase the rate to $8.00—which was voted in February 1948.

The Biometrics Section of the ASA participated actively in the International Conferences and in the organization of The Biometric Society, the new international organization of biometricians. The publication *Biometrics* completed its first year as a quarterly and was requested to print the Proceedings of the First International Biometric Conference.

The Temporary Committee on the Training of Statisticians, under the chairmanship of Francis G. Cornell, made a report to the Board. As a result of this report, a Section on the Training of Statisticians was organized within the ASA and is charged with the formulation of standards for teaching statistics and with the preparation of special sessions at the Annual Meeting in 1948.

The Association has been actively involved in preparations for the 17th Decennial Census of the United States. The Committee to Advise the Bureau of the Census, which has the responsibility of reviewing all requests from the general public for the inclusion of additional information in census data, has been working as a clearing house between statisticians and government to determine the content of census questionnaires.

The Advisory Committee to the Bureau of Labor Statistics on the Employment Statistics Program submitted a comprehensive and final report in 1947 which was summarized in the April 1948 issue of *The American Statistician.*

The year has been one in which the Association has grown, its responsibilities have increased and the services it renders to the statistical profession have expanded. The influence of ASA publications has been notably enhanced. Thus, the circulation of the *Journal* gained by approximately 20 per cent to reach 5,895. *The American Statistician* is reaching a broad group of less technically and mathematically minded readers, familiarizing them with general developments in the statistical field. In recognition of the probability that the Association must in the coming years provide still broader services, a decision has been taken to launch a new publication, under the editorship of Oscar K. Buros, probably in 1949. This new organ will cover the field of book reviews, will provide an index of statistical methodology, and will make available a comprehensive digest of all significant articles, both domestic and foreign, which deal with statistical problems. The increasing reliance of science, industry, business and government on statistical method, coupled with the continuing development of the science of statistics itself, confronts your Board with the problems incident to endeavoring to meet the unsatisfied needs of a corps of workers in an expanding science.

LESTER S. KELLOGG
ISADOR LUBIN
LOWELL J. REED
WALTER A. SHEWHART
SAMUEL A. STOUFFER
WILLARD L. THORP
HELEN M. WALKER
SAMUEL S. WILKS

The Secretary's Report on Membership

During 1947 membership in the American Statistical Association increased by 1,028, bringing the total on December 31, 1947 to 4,871. Three new corporate members joined the Association. The continued growth during the past year indicates the growing interest in the Association's activities and a large potential membership.

The 1947 membership is composed of the following groups:

Honorary members	15
Fellows	142
Regular members	4,714
Total Membership	4,871
Corporate Members	6

There were 82 resignations in the course of the year, and 189 members were removed from the rolls for failure to pay dues.

The members of the Biometrics Section, at the end of December, 1947, numbered 1,060, of which 262 were associate members and 798 were also members of the ASA.

The deaths of the following members were recorded during the year: Irving Fisher, *Fellow;* Charles A. Bauer, L. A. Beers, H. M. Biggert, Raymond Carpenter, Charles Cresswell, Kenneth Duncan, Albert M. Freeman, Thomas B. Graham, Otto Jeidels, James E. Kelley, Sara A. Kerr, H. G. MacCall, Joseph P. McElroy, Hans Nybolle, Henry A. Ruger, James G. Smith, Philip Sweet, Henry A. Van Pelt, *Regular Members.*

LESTER S. KELLOGG, *Secretary*

Report of the Treasurer

The 1947 budget as approved by the Board of Directors planned for an income of $32,535.00 and expenditure of $35,192.00. At the end of the year actual income was $37,014.87, and our expenditures came to $42,134.72.

The increase in income has come largely from new memberships, increases in subscriptions and additional advertising in the *Journal*, and was substantial enough to permit the publication of a greatly enlarged December issue of the *Journal.* Income from *Biometrics* has increased, and it is expected that its income for the next few years will continue to expand as its service becomes more widely recognized.

The national office is planning a new drive for memberships and subscriptions this year, and the budget will consequently be substantially increased.

LESTER S. KELLOGG, *Treasurer*

Report of the Auditors

To the Board of Directors of
American Statistical Association

We have examined the attached financial statements of American Statistical Association relating to the year ended December 31, 1947. Our examination was made in accordance with generally accepted auditing standards, and included such tests of the accounting records and other supporting evidence and such other procedures as we considered necessary in the circumstances.

The recorded cash receipts for the year were traced to the deposits shown on the bank statements and the amounts for dues and subscriptions were tested with the membership and subscription records. The paid checks were inspected and related vouchers tested in support of cash disbursements for the year. The bank balances were reconciled with amounts reported direct to us by the depositaries and the cash on hand and the securities owned at December 31, 1947 were verified by inspection. We did not check the membership and subscription records in detail or make any independent verification of the inventory of old Journals, the office records of which are based, in part, on data assembled in prior years, no recent physical inventory having been taken.

The life membership reserve is computed on the basis of the combined annuity table of mortality with assumed interest at 4% per annum and an assumed annuity of $5.00 per life member, in accordance with a resolution of the Board of Directors on March 31, 1936. The amount treated as income in each year represents the excess of the reserve at the beginning of the year plus interest for the year and new life membership receipts over the required reserve at the end of the year.

In our opinion, the accompanying statements present fairly the position of American Statistical Association at December 31, 1947 and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent in all material respects with that of the preceding year.

PRICE, WATERHOUSE & CO.

AMERICAN STATISTICAL ASSOCIATION

BALANCE SHEET

	December 31,	
Assets	*1947*	*1946*
Cash in bank and on hand	$ 1,853.91	$12,475.91
Accounts receivable	1,356.22	774.19
Investments:		
United States Savings Bonds, Series D, at redemption value	6,138.00	5,910.00
Stocks, at cost (at market quotations $6,150 and $6,358, respectively)	5,793.50	5,793.50
Inventory of old Journals, at approximate cost	2,024.63	2,158.04
Furniture and equipment, at cost less depreciation	2,482.76	847.11
	$19,649.02	$27,958.75
Liabilities		
Accounts payable	$ 6,515.76	$ 3,537.45
Deferred income (collections applicable to subsequent year):		
Dues	$ 1,054.50	$ 7,260.67
Subscriptions	2,704.76	3,158.48
Other	115.35	
	$ 3,874.61	$10,419.15
Net worth:		
Life membership reserve	$ 3,750.20	$ 3,377.35
Centenary Sustaining Fund		6,376.01
Surplus, per statement	5,508.45	4,248.79
	$ 9,258.65	$14,002.15
	$19,649.02	$ 27,958.75

AMERICAN STATISTICAL ASSOCIATION

SURPLUS STATEMENT

	Year ending December 31,	
	1947	*1946*
Balance at beginning of year	$ 4,248.79	$ 4,594.20
Add—Transfer of balance in Centenary Sustaining Fund	6,379.51	
	$10,628.30	$ 4,594.20
Deduct—Excess of expenses over income for the year, per income statement	5,119.85	345.41
Balance at end of year	$ 5,508.45	$ 4,248.79

AMERICAN STATISTICAL ASSOCIATION
INCOME STATEMENT

	Year ending December 31, 1947	1946
Income:		
Dues—current year	$21,298.58	$17,452.50
Dues—prior years	215.00	92.00
Life memberships	397.33	352.47
Subscriptions	6,206.95	5,341.30
Advertising	2,010.08	1,700.88
Reprints	509.00	530.93
JOURNAL sales	2,827.80	793.80
Biometrics Section income	2,466.86	2,121.97
Miscellaneous	684.70	394.71
Dividends and interest (after deducting $135.09 in 1947 apportioned to life membership reserve)	398.57	432.08
	$37,014.87	$29,212.64
Expenses:		
JOURNAL—printing, mailing and reprints	$10,105.81	$ 7,908.35
Bulletin	4,259.88	991.01
Salaries and wages	17,037.26	11,876.44
Rent	1,365.00	1,020.00
Office supplies, printing and mimeographing	1,764.42	1,277.44
Postage	1,511.11	1,060.07
Telephone and telegraph	512.03	369.30
Travel expense—officers	603.91	287.29
Biometrics Section expenses	2,366.46	2,185.81
Membership Directory expense		1,575.76
Storage of old JOURNALS	72.00	72.00
Cost of old JOURNALS sold	382.27	58.69
Depreciation of furniture and equipment	324.99	149.49
Annual meeting expense (net)	647.89	238.84
Loss on disposal of equipment	151.09	
Miscellaneous	1,030.60	487.56
	$42,134.72	$29,558.05
Balance, loss, carried to surplus	$(5,119.85)	$ (345.41)

BOOK REVIEWS

Edited by
Oscar Krisen Buros
Rutgers University

Statistical Methods in Research and Production: With Special Reference to the Chemical Industry. Edited by *Owen L. Davies*. Written by *Lewis R. Connor*, *Wilfred R. Cousins*, *Owen L. Davies*, *Francis R. Himsworth*, *Harold Kenney*, *Wilfred L. Stevens* and *Charles C. Tanner*. (Chemists, Engineers, and Statisticians employed by Imperial Chemical Industries Ltd., Manchester, England.) Published for the Imperial Chemical Industries Limited. Edinburgh 1, Scotland: Oliver and Boyd Ltd. (Tweeddale Court), 1947. Pp. xii, 292. 28*s*.

Review by D. B. DeLury
epartment of Mathematical Statistics
and
A. E. R. Westman
Director, Department of Chemistry
Ontario Research Foundation, Toronto 5, Canada

The object of the authors is to put into the hands of the industrial research worker a handbook prepared especially for him in terms with which he is likely to be familiar. To this end, the main body of the discussion is given in verbal and arithmetical terms. Whatever algebraic developments are needed appear in appendices at the ends of chapters. This is not to imply that the book is chiefly a guide to effective computational methods.Emphasis is placed on the experimental situation and much careful and thoughtful explanation accompanies each example.

The range of topics covered is indicated by the chapter headings: Frequency Distributions, Averages and Measures of Dispersion, Tests of Significance, Analysis of Variance, Regression and Correlation, Frequency Data and Contingency Tables, Sampling, Control Charts, Prediction and Specification.

Each chapter starts with a general description of the questions with which it deals, followed by examples based on data (usually drawn from the chemical industry) which are intended to illustrate the type of situation with which the chapter deals. The discussion of the examples is not obscured by arithmetical details; systematic methods of computation are dealt with in chapter appendices.

This plan is, no doubt, effective, but starting as it does with the data rather than with the *gathering* of the data it tends to cover up the importance of statistical considerations in the planning of the experiment. Perhaps each chapter should include an example in which statistical techniques can extract no worth-while information from the data!

Those who are interested chiefly in practical applications will find much useful information gathered together in this handbook. On the whole it is clearly presented and illustrated by good examples. The reader who has a knowledge of theoretical statistics will doubtless find many points on which he does not agree with the authors and others may find some of the discussions confusing. Certainly the quality of the book is uneven with respect to both clarity of exposition and accuracy of statement. While many of the discussions and explanations are excellent, others are obscure, sometimes inaccurate and occasionally meaningless. For example, it is not likely that anyone will be helped to a clearer understanding of statistical estimation by the statement that "Statistical estimation consists, in general, of postulating that a given set of data or *sample* was drawn from a Universe of a certain form. The constants or *parameters* required to define the Universe may be estimated from the data or may be postulated" (p. 14).

Owing, perhaps, to reluctance of the authors to expand the book to greater length, some topics are treated too briefly. Only one model for the analysis of variance is explained, that in which differences among cell averages are regarded as arising entirely from random sources. The case of "fixed effects" is not discussed, although these methods are used to exhibit, for example, differences among three "gaugers," who could hardly be considered a randomly chosen sample from a population of gaugers, in a test of the compressive strength of Portland cement (Example 5.5), and the difference between two operators who carried out simultaneous determinations of the percentage of ammonia in a plant gas (Appendix 4B).

The chapter on regression suffers considerably from the failure to discuss the underlying statistical model. The conditional nature of the regression problem never emerges. Instead, there is created a hierarchy of regression problems, the two extremes of which are, apparently, (*i*) the case in which the quantities observed are inherently variable, in a statistical sense; (*ii*) the case in which the quantities observed are not statistical variables; the only random variation in this case comes from errors of measurement. This is described as the case in which the variables are related by an exact mathematical formula.

The choice of examples to typify these two extremes is unfortunate. Case (*i*) is discussed in terms of the relationship between hardness and abrasion loss in rubber. The observations on these two variables are made on samples of rubber in which the amounts and types of ingredients are deliberately changed from one sample to another, a situation which would seem to require the use of multiple regression methods. Case (*ii*) is illustrated by the dependence on concentration of the response to a penicillin solution. It may be granted that the concentrations are known without appreciable error, but the variations in response represent genuine biological variation and it is difficult to envisage an "exact" relationship between response and concentration. In any case, the differences between these two cases largely disappear when the true nature of regression is recognized.

The distinction between these types is maintained throughout, culminating in two different formulae for the standard error of a prediction based on a regression equation (Formulae 6.3 and 6.31). It is not clear what is being predicted, a mean value or the value of a future observation, but whichever is intended, the same formula (6.3 *or* 6.31) should apply in both case (*i*) and case (*ii*).

More serious is the failure to bring together regression methods and the analysis of variance, so as to present the elements of the analysis of covariance, a technique second in importance only to the analysis of variance in all fields of experimentation.

Inevitably a few slips must creep into every book. The following may be noted.

(1) The F test is introduced in a comparison of variance estimates from two independent samples, a situation in which a two-tail test usually is indicated. However, the test employed is one-tailed (Example 4.51).

(2) Throughout the discussion of the errors of prediction which result from the errors of the regression coefficients (p. 130), a should be replaced by $\bar{y}$.

Readers interested in control chart techniques will find a clearly written chapter on the subject based on British Standard practice.[1] However, only the charts for average and range are discussed in detail and the unfortunate impression is given by some parts of the discussion and by the numerical example that a subgroup of four or five is necessarily "rational." In the chapters concerned with other techniques, not much attention is paid to the quality control approach to industrial problems and readers familiar with developments in America will consider the following statement on page 4 of the Introduction to be overcautious: "Its application in the wider field of the chemical industry is still a matter of some doubt, but in view of its striking success in other fields, it is felt that the possibility of its application to chemical plants should receive careful consideration."

The book, as it stands, should be most useful to the industrial investigator. The intention of the authors, stated in the preface, to keep the handbook up to date, will furnish an opportunity for the critical editorial overhauling which is needed to bring this work to a uniformly high level of excellence. A sequel on the planning of experiments is contemplated.

The typography is exceptionally good. Misprints are rare and unimportant.

REVIEW BY E. L. LEHMANN

Assistant Professor of Mathematics
University of California, Berkeley, California

AUTHORS of any book such as the present manual who, without assuming any previous statistical knowledge on the part of their readers, want to give an introduction to intelligent use of statistical methods are concerned with three different aspects of statistics. (*a*) They must explain the basic concepts such as probability, expectation, statistical hypothesis, power of a test, confidence intervals, etc. (*b*) They must describe the most important statistical procedures, state the conditions under which these procedures are known to be valid, and discuss some of the computational problems that arise. (*c*) They must illustrate the methods with representative examples from the field of application with which they are concerned.

---

[1] Dudding, B P., and Jennett, W. J. *Quality Control Charts.* British Standard 600R: 1942. London: British Standards Institution, 1942. P. 89.

In the present volume, as in most other books of this kind, (*b*) and (*c*) are emphasized at the expense of (*a*) and even in the discussion of particular procedures the underlying assumptions are frequently indicated only rather vaguely. This seems regrettable since, as Professor Wolfowitz pointed out in a recent review in this JOURNAL (41: 127–8 Mr '46), "misapplication of statistical methods is more likely to be serious than the use of inefficient computing techniques." In the discussion of hypothesis testing, for example, no mention is made of possible alternatives or, therefore, of power. As a result the assumptions underlying a test are never quite clear. For instance, when the authors explain their recommendation of a two-tailed F-test in analysis of variance, the class of alternatives which they have in mind seems to change during the discussion.

The material covered in the book is fairly conventional. Two introductory chapters on frequency distributions, and averages and measures of dispersion are followed by a discussion of tests of significance (t-test and F-test—the authors make the common slip of using a one-tailed F-table for a two-tailed test without adjusting the level of significance). Then come the two main chapters on analysis of variance, and regression and correlation, and chapters on frequency data and contingency tables, sampling, and control charts. At the end are a few pages on prediction and specification, which are mainly concerned with the Camp-Meidell and Tchebycheff inequalities.

There is a somewhat incomplete glossary of statistical terms and an unusually comprehensive collection of tables (area under normal curve, percentage points of χ^2, t and F; values of correlation coefficient for different levels of significance; limits of the expectation for binomial and Poisson distributions; and control chart limits for average and range). A nice feature of the book is that the methods are illustrated throughout with many interesting examples from the chemical industry.

The authors state that they intend to keep the handbook up to date and indicate that they hope to include some sequential techniques in a later edition. It would perhaps be useful also to include some nonparametric methods.

Selected Techniques of Statistical Analysis: For Scientific and Industrial Research and Production and Management Engineering. By the Statistical Research Group, Columbia University, Applied Mathematics Panel, Office of Scientific Research and Development. Edited by *Churchill Eisenhart* (Chief, Statistical Engineering Laboratory, National Applied Mathematics Laboratories, National Bureau of Standards, Washington, D. C.); *Millard W. Hastay* (Research Staff, National Bureau of Economic Research, New York 23, N. Y.); and *W. Allen Wallis* (Professor of Statistics and Business Economics, University of

Chicago, Chicago 37, Ill.). New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1947. Pp. xiv, 473. $6.00. (London W.C. 2: McGraw-Hill Publishing Co., Ltd. [Aldwych House, Aldwych], 1948. 36s.)

REVIEW BY HENRY SCHEFFÉ
Associate Professor of Engineering, University of California
Los Angeles, California

THIS book consists of 17 quite unrelated chapters dealing with various problems which arose in projects undertaken by SRGC. Almost anyone confronted with practical problems calling for statistical methods may be assured of finding in this book material of value to him, providing he has some background of statistical knowledge, perhaps the amount acquired in two year courses in statistics. A very thorough index of 43 pages of fine print increases the usefulness of the book. It would not be suitable for a textbook for various reasons, including the specialization of the subject matter.

Since a summary of the book by chapters would take too much room here we shall list the titles and number of pages for each chapter, and add a few remarks. Part I, Industrial Statistics, consists of chapters 1 through 6; Part II, Planning Experiments, of chapters 7 through 13; and Part III, Techniques and Tables, of chapters 14 through 17. The chapters are:

1. Use of variables in acceptance inspection for per cent defective (92 pp.).
2. Tolerance limits for normal distributions (16 pp.).
3. Multivariate quality control, illustrated by the air testing of sample bombsights (74 pp.).
4. Effects of rounding or grouping data (40 pp.).
5. Acceptance inspection when lot quality is measured by the range (10 pp.).
6. Variability of amount of inspection for double, multiple, and sequential sampling (10 pp.).
7. Planning and analyzing experiments for comparing two percentages (20 pp.).
8. Planning and interpreting experiments for comparing two standard deviations (52 pp.).
9. Utilization of limited experimental facilities when the cost of each measurement depends on its magnitude (10 pp.).
10. Experiments in which the independent variable is known only in terms of a probability distribution (10 pp.).
11. Planning an experiment for estimating the mean and standard deviation of a normal distribution from observations on the cumulative distribution (14 pp.).
12. Abandoning an experiment prior to completion (10 pp.).
13. Planning experiments seeking maxima (10 pp.).
14. Probability that sample means are in opposite order to population means (8 pp.).
15. Significance of the largest of a set of sample estimates of variance (12 pp.).
16. Inverse sine transformation of proportions (22 pp.).
17. Elements of sequential analysis (13 pp.).

The chapters were written by Albert H. Bowker, Churchill Eisenhart, Milton Friedman, Millard W. Hastay, Harold Hotelling, Edward Paulson, L. J. Savage, Herbert Solomon, and W. Allen Wallis. Some of the chapters are based on work credited to the following people not included in the above list of writers: Kenneth J. Arnold, M. A. Girshick, Myra Levine, A. Wald, and J. Wolfowitz.

Among the nonuniform aspects of the chapters is the mathematical preparation assumed on the part of the reader: At one extreme we find 7 pages devoted to computational methods for the evaluation of $\bar{x}+ks$, where $\bar{x}$ is the sample mean, s, the sample standard deviation, and k, a given constant (this may be justified by the resistance to the use of s in certain industrial quarters), at the other extreme there is mention of a set of measure zero (in a "technical note"). Mostly the authors have conscientiously tried to meet the needs of the "practical man" for whom the book was written, and if he happens to be also a mathematician he may occasionally find somewhat painful* the verbal arguments where he would use symbolic. If unifying motifs are sought in this work, one is to be found in the wide use of the concept of the power function or operating characteristic of a statistical test; this occurs in chapters 1, 5, 6, 7, 8, 12, 14, and 17. The reviewer considers one of the main virtues of the book to be the influence it will have in making this concept a tool of the practitioner.

The book brings to the reader a large amount of material which has been available before only in periodicals. Not even thus available has been much of the content of chapter 3 on multivariate quality control. Aside from the novelty of applying the T^2-type of statistic to quality control problems, we note also the following innovations: The distance apart of the control lines on the control charts depends on the length of the base period from which they are calculated (the distance decreasing with increasing length). "Yardsticks" are introduced on the control charts which allow us to judge whether the difference of the ordinates of two points is statistically significant. Where the ratio rather than the difference is statistically appropriate, the yardstick device is retained by use of a logarithmic vertical scale.

It is clear that this group did some very fine statistical work. For their efforts in rescuing some of the results of permanent value from burial in the files the director of the group, Professor Wallis, and the other editors deserve congratulations.

The Design of Experiments, Fourth Edition, *Ronald A. Fisher* (Arthur Balfour Professor, University of Cambridge, Cambridge, England). Edinburgh 1, Scotland: Oliver and Boyd Ltd. (Tweeddale Court), 1947. Pp. xi, 240. 12*s.* 6*d.*

REVIEW BY A. E. BRANDT
Statistical Consultant, Office of Technical Director
Naval Ordnance Laboratory, Washington, D. C.

THE reviewer has enjoyed the opportunity to appraise the fourth edition of *The Design of Experiments* after a number of years of teaching and applying the principles of design presented. These principles were first received by the reviewer in a course of lectures previous to publication of the first edition. Though each previous edition has been read and studied, the reviewer was well rewarded for the time spent on the latest one.

---

* Also that the one formally stated theorem in the book (p. 409) is possibly false, the proof certainly is, and in the two applications to discrete distributions the conditions of the theorem are obviously not satisfied [condition (14)]. Among the errors in the proof is the unjustified equating of $E(XY)$ to $E(X)E(Y)$ in several places. A correct theorem of the sort attempted, namely, one that is easily and widely applicable, would be nice to have.

Familiarity with the statistical procedures presented by the author in his *Statistical Methods for Research Workers* is taken for granted as a background for understanding this book. However, the pleasing literary style and nonmathematical treatment in all but the last chapter have led some persons who lack this background to expect great benefits from reading it. Those readers who complain that the book is hard to understand are admitting their own shortcomings.

The course of lectures given by the author previous to the publication of the first edition of this book was called *The Logic of Experimentation*. The first two chapters and, to a large extent, the third, are more aptly described by this title than by the title of the book. In the first chapter the author presents a convincing argument for his theory of inductive inference and his reasons for rejecting inverse probability. By means of his classic experiment relating to a lady's claim that she can determine by taste whether the milk or the tea has been placed in a cup first, the whole logical structure of critical experimentation, from the statement of the question to the conclusion to be drawn from the experimental results, is developed. The roles of randomization and repetition or replication and the necessity of their inclusion in a critical experiment are made clear.

The data on Zea mays from Darwin's historical experiment on growth rate and Galton's interpretation of them are used to show how fallacious uses of statistics lead to discordant results and as a warning against the manipulation of data. The effect of local control, such as pairing or grouping, on the sensitiveness of each individual observation is clearly shown. The general test of the hypothesis that two samples are random samples from identical populations presented in the last section of the third chapter merits greater study by statisticians who lack experience in experimentation in some field of science such as biology, medicine, or physics.

To give a clear understanding of the logic of experimentation presented in the first three chapters and its practical application, the author uses a number of examples in chapters 4 to 9 inclusive. In the fourth and fifth chapters, simple situations are presented which can be handled by means of randomized block or latin and higher square designs which are the simple or basic types. In the next four chapters, designs suitable for situations of greater and greater complexity are illustrated. The advantages of the factorial arrangement over a series of experiments in which but one factor is varied at a time are emphasized in chapter 6. The use of local controls such as confounding and partial confounding and of statistical control through concomitant measurements are dealt with rather fully in chapters 7, 8, and 9.

Chapter 10 on the generalization of the null hypothesis and fiducial probability contains much of interest to the practical experimenter but is especially important for theoretical statisticians. The final chapter on the measurement of information is especially commended for serious study, not because of the calculus involved, but because the principles in it are ignored so frequently.

Though the author has used the results of actual experiments to illustrate the requirements of sound and intelligible experimental designs and means of drawing valid conclusions from them, he has avoided producing a compendium of designs which have been used successfully in various fields of experimentation. It would seem, however, that there is a distinct need for such a book.

Das Harmoniegesetz der Statistik: Eine Untersuchung über die metrische Interdependenz der sozialen Erscheinungen. [The Law of Harmony in Statistics: An Investigation of the Metrical Interdependence of Social Phenomena.] *L. V. Furlan* (Professor at the University of Basel). Basel, Switzerland: Verlag fur Recht und Gesellschaft A.-G., 1946. Pp. xii, 504.

REVIEW BY HILDA GEIRINGER
Professor of Mathematics, Wheaton College
Norton, Massachusetts

THIS book presents an investigation on statistics rather than a statistical investigation. As the author asserts, the starting point is basically new. The purpose is to study great masses of statistical numbers without any consideration of their meaning but with respect to some "number-theoretical" properties. The result is a very definite statement concerning a very vaguely defined subject: Consider a mass of statistical numbers, "the way they are found in every statistical yearbook." They form together some kind of reflection of the social life of a group of people, of a nation, or of groups of nations. Such a mass of statistical numbers, no matter whether it consists of only a few or of very many numbers, represents a sample of the infinite mass of numbers which describe the social phenomenon, or rather some aspects of the social phenomenon. As an example a "collective" consisting of 1.19×10^4 numbers is considered (quoted from *Statist. Handbuch der Weltwirtschaft,* Berlin, 1936), containing data on various fields of economical life for eighty countries for the year 1933. (This collective is analyzed as a whole and also resolved into 80 partial collectives.) Another example contains an even larger mass of numbers, practically all numbers contained in the *Statist. Jahrbuch der Schweitz, 1912.* (In each of these and other "natural" collectives certain adjustments are made before the analysis of the material is started.)

The main problem and aim of this analysis is a very strange one. The first figure in a natural number is called its "indicant." If decimal fractions are considered, the first nonzero number is the "indicant." In a decadic system there are nine possible indicants. For each of these nine indicants, their relative frequency within the considered collective of numbers is computed.

The author's contention is that, quite independently of the meaning of these statistical numbers, if the "natural collective" is only large enough and somehow a random collective, there exists a striking regularity concerning these nine relative frequencies. (These nine frequencies form the "spectrum" of the natural collective.) The statement is that for a natural collective the relative frequency of the indicant i, ($i=1, 2, \ldots, 9$) equals $\log(1+1/i)$. If this is so, i.e., if the relative frequency of the indicant (i) has this value, the spectrum is called the "harmonic spectrum," and this validity of the harmonic spectrum for natural collectives constitutes *the harmony law of statistics*. A particular consequence of this law is the fact that the indicant "one" appears with relative frequency $\log 2 = .30$; hence, it appears much more frequently than it would in a uniform distribution of the nine numbers. This is called "concentration on the one" and is considered as the most striking and the most important feature of the law. The frequency of the numbers $1, 2, \ldots, 9$ then decreases according to this law. This "harmony law" which forms the main result of the 500-page work may be stated once more: "The spectrum characteristic of natural collectives is the harmonic spectrum" (p. 443). Or, in the more specific form: "The indicant of a statistical number is a chance variable which assumes the values $i=1, 2, \ldots, 9$ with probability $\log(1+1/i)$" (p. 468).

As a theoretical background for these statements certain "artificial," i.e., mathematically defined, collectives (sequences) are considered with respect to their "spectra." Examples are: (*a*) the sequence of natural numbers $\{n\}$ (where $n=1, 2, \cdots$); (*b*) the sequence of the squares of natural numbers $\{n^2\}$ ($n=1, 2, \cdots$), of the cubes, etc.; (*c*) the sequence a^n (where a is fixed and $(n=1, 2, \cdots)$; (*d*) the sequence of all rational numbers in the well known order $\frac{1}{n}, \frac{2}{n-1}, \frac{3}{n-2}, \cdots, \frac{n}{1}$ ($n=1, 2, \cdots$) (fractions with a common divisor of numerator and denominator are discarded). For any such sequence the relative frequency of the indicant i within the first n elements is computed, and the limit of this relative frequency as $n \to \infty$ is considered. It is easily seen that for the first two examples such a limit does not exist. (Such "collectives" are termed "erratic.") On the other hand it is shown that for (*d*), the limit distribution of the indicants is the harmonic distribution; i.e., the relative frequency of the indicant i equals $\log\left(1+\frac{1}{i}\right)$.

It is not the reviewer's business to express an opinion as to whether such facts are of mathematical interest from the point of view of pure mathematics. In the reviewer's opinion the main problem consists obviously in constructing a connection between them and the harmonic property of natural collectives, a connection which would allow consideration of a certain theoretical scheme as a "model" of a natural collective. Such a "model" is, in fact, indicated, namely the "artificial collective" (*c*): $\{a^n\}$

($n=1, 2, \cdots$), which shows "approximately" the harmonic spectrum (p. 303). "The exponential function which expresses the law of natural growth forms a scheme of the many small causes which act by way of multiplication, being unobservable as long as they are isolated but becoming manifest by means of the spectra of natural collectives" (p. 503, translated and condensed). "In this way the natural collective becomes an instrument which reveals certain properties of the social phenomenon which could not have been found in any other way" (p. 502). Thus, on the basis of the interpretation of natural collectives by means of the sequence $\{a^n\}$ a theory of the natural collectives is somehow attempted. This theory is, to say the least, extremely vague.

It may well be that a mathematician is not in the position to do justice to a work belonging to the field of social sciences. With this reservation a few comments may be allowed. It is hard to decide whether it is possible to give precise definitions to the main concepts introduced in the statement of the harmony law, as "natural collective," "statistical number," etc. There seems to be little point in making mathematically precise statements concerning incompletely and vaguely defined subjects. If such definitions are given the questions will remain: (*a*) whether this law expresses a theoretically consistent statement, and (*b*) whether in its application to reality it states a significant feature of reality, i.e. of the social reality behind the natural collectives. This seems to the reviewer rather improbable, certainly so far as mathematically precise statements are concerned.

If we look in a purely empirical way at the examples in Furlan's book (e.g., p. 22) or if we examine, on our own, some statistics in yearbooks, it actually seemed to the reviewer as if there existed a certain vague "concentration on the one" as first figure. But here again some rather obvious remarks force themselves upon us, e.g., that this concentration may be destroyed by a change of unit ("100 square miles" instead of "100 hectares," or "per 5000 people" instead of "per 1000 people"), or that statistical numbers have upper and lower bounds (think of the possible numbers of children in a family, to quote a trivial example) which influence the indicants. If we consider this question of unit a little closer it seems that the "harmony law" ought to be independent against multiplication by a constant in order to be independent of the unit of measurement. This would imply the following theoretical statement: Consider an "artificial collective," i.e., a sequence of numbers where the indicant, i, appears with a limit of relative frequency equal to $\log(i+1)-\log i$, as postulated by the harmony law. If in such a sequence each number is multiplied by one and the same constant the relative frequency of the nine indicants should remain unaltered. This however is certainly not true. As a simple conterexample consider a sequence consisting only of the nine numbers, $1, 2, \cdots, 9$ which are each written down so often that the number i occurs with the frequency (or more precisely with the limit of relative frequency) equal to $\log(1+1/i)$. Now form a new sequence by multiplying each number in the original sequence by the constant 2. This new sequence will consist of the numbers 2, 4, 6, · · · , 18 where the "2" will appear with frequency log 2, the "4" with frequency log $\frac{1}{2}$, · · · the "18" with frequency log 10/9. In this new sequence the indicant "1" appears in 10, 12, 14, 16, 18 with frequency equal to $\log 6/5+\log 7/6+\cdots+\log 10/9=\log(6/5\cdot 7/6\cdot 8/7\cdot 9/8\cdot 10/9)=\log 2$, hence in fact again with frequency log 2. But, e.g., the indicants 3, and 5, and 7, and 9 will each have frequency zero instead of the frequencies postulated by the harmony law which thus appears to be completely destroyed by this operation.* Undoubtedly such obvious remarks must have occurred to the author. They apply, in this form, of course to artificial, i.e., mathematical collectives only. A certain empirical regularity like the concentration on the one may be observable on "natural collectives" under certain specified conditions. However, an empirical observation, even if correct and fairly general hardly constitutes a convincing basis for a far reaching general mathematical statement. And the theoretical foundation by means of a mathematical model (see our remarks concerning $\{a^n\}$ can hardly be considered at this stage as much more than an aperçu.

An impeding feature of the book is the lack of an index. This is particularly disagreeable since so many new and unusual words and notations are used. Otherwise the work is certainly extremely well written and very original. It remains to be seen whether this whole problem of the relative frequency of the indicants of statistical numbers reveals—by means of a suitable interpretation—aspects of social phenomena which are of interest to social scientists.

Fundamentals of Statistics. *Truman Lee Kelley* (Professor of Education, Harvard University, Cambridge, Mass.). Cambridge 38, Mass.: Harvard University Press (38 Quincy St.), 1947. Pp. xvi, 755. Lithotyped. $10.00. (London E.C. 4: Oxford University Press [Amen House, Warwick Square], 1948. 55*s*.)

REVIEW BY CYRIL BURT
Professor of Psychology, University College
University of London, London, England

THIS book is more than a revised and expanded edition of Kelley's well-known *Statistical Method* (Macmillan Co., 1923). But it preserves many of the admirable features of that excellent work. In his approach, as the author pointed out at the time, his "greatest inspiration" had come from Karl Pearson; and he found "the English School had been most contributive." The present volume shows how Pearson's general conception of statistics and the numerous contributions that he made on fundamental problems may still serve as the foundation on which the newer additions of more recent years may be superposed.

---

* It is obvious that the same conterexample holds if instead of the logarithmic frequencies, postulated by the harmony law, any other nine frequencies, $p_1, \ldots p_9$ with sum equal to one, are assumed. It seems to the reviewer that from a purely mathematical point of view, it is, the concept of the "indicant" which is open to objections rather than the particular frequency distribution considered by the author.

The chief alteration in the new volume lies in the greater attention paid to the logical principles underlying statistical reasoning. Statistical procedures are presented, not as a "bag of useful tricks and dodges," but as a mode of reasoning about complex problems. This point of view is developed clearly and convincingly in the opening chapter.

The chapters that follow, like those of the earlier book, deal in order with statistical series and statistical tables, graphic methods in two and three dimensions, and the "stable features of phenomena" (a topic that was only briefly touched upon in the earlier volume). This leads to the methods of measuring central tendency and variability, and the properties of the normal distribution. A fresh chapter is then inserted on the statistics of attributes; the use of chi-square and the variance ratio is introduced in this connection. We then proceed, much as before, to the problems of correlation, regression, and multiple correlation. The latter, it may be remembered, Kelley dealt with on Pearsonian lines in terms of determinants. It is therefore natural for him now to introduce the uses of matrix methods in this connection.

The original chapter on "sundry statistical problems and procedures" is considerably enlarged. Some of the newer topics are those of periodicity, lead and lag in time series, direct and inverse interpolation, the variance errors of different statistics, and (another welcome addition) sequential analysis.

The last chapter is entitled "mathematics: the mentor of statistical ingenuity." Its main purpose is to provide, in compact form, a mathematical background for the experimentalist, so that he can equip himself with a knowledge of the particular functions that are most likely to be suggestive of phenomenological relations. Accordingly, in addition to a discussion of the point binomial, the Poisson distribution, and the hypergeometric series, we have sections on the Fourier series, factorials, growth curves, matrix algebra, the determinantal solution of simultaneous equations, and other topics. The book concludes with a useful set of statistical tables, a mathematical background test, and lists of symbols and common mathematical terms (with definitions), and a key to formulae.

It will be seen that, while the book includes nearly all the stock topics, it also contains much useful matter that is not to be found in the usual type of statistical textbook. As the author points out, statistics has developed so rapidly during the last twenty-five years that it would be impossible for anything short of an encyclopaedia to cover the entire ground. It is therefore easy for the critic to point out omissions. The psychologist will probably regret that more attention has not been given to the analysis of variance and to factor analysis. In the discussion of the logical basis it is perhaps disappointing that more space has not been given to the theory of probability and its relevance to tests of significance, and particularly to the latest developments in the theory of estimation.

REVIEW BY HORACE W. NORTON
*U. S. Atomic Energy Commission**
Oak Ridge, Tennessee

FUNDAMENTALS *of Statistics* is an elementary textbook written by a well-known psychologist. Although the early chapters usually employ little mathematics, the later chapters use a fair amount of algebra (especially in treating determinants and matrices), and there is an occasional appeal to the calculus. The book is addressed to no specified class of readers, but it is fair to say that in flavor, language, examples, references, and point of view, it is definitely of the social sciences. Its object is clearly stated in the following quotation:

> The usual elementary text in statistics seeks to acquaint students with the techniques necessary to obtain . . . averages, measures of variability [etc.]. . . . This text considers these things as a secondary aim, the primary aim being to acquaint the student with the logic back of these procedures. It is believed to be more important . . . to understand why a mean is employed than to know the steps for calculating it . . . (p. 57).

As a text, it should be regarded as of more nearly 500 pages than 750, roughly 200 pages being given to miscellaneous aids, such as mathematics, references, tables, and indexes. There are a few problems, and much emphasis on vocabulary.

The first two chapters are introductory, exhibiting to the reader the meaning, the scope, and the uses of statistics. Thirty pages are given to a discussion of the mathematical background needed for statistical reading, quoting the Social Science Research Council report (1932) and including brief explanation or comment on a number of elementary mathematical topics. This is unusual by itself, and it is therefore especially noteworthy that there is also a mathematical background test with instructions and scoring key, enabling the student to assess objectively his readiness to read this book. The next nine chapters (275 pages) are conventional in subject matter, treating statistical tables, graphics, the notion of stable features of variable phenomena, measures of variability and of central tendency, the normal distribution, attributes, regression and correlation, and multiple regression. Chapter 13 is a miscellany, including time series and periodicity, methods of arithmetic such as machine extraction of roots, and sequential analysis. Chapter 14 is chiefly mathematics, but some statistical topics occur, such as moments of the binomial and Poisson distributions, and some transformations. The fifteenth and last chapter gives a few mathematical and statistical tables and a considerable list of references to such tables. Appendix A gives the mathematical background test already mentioned. Appendix B lists symbols and terms, formulas given in the text, and references.

Because the author has not designated a specific group of readers as his target, critical evaluation is not easy. Moments are emphasized, and there is a lengthy discussion of the correlation coefficient, its deficiencies, corrections, interpretation, and uses; and of matrices and determinants; solution of linear equations; and partial and multiple regression. The analysis of variance is mentioned a number of times but without systematic treatment. There is nothing on independent comparisons or single degrees of freedom, no mention of components of variance, no analysis of variance tables, and no worked example. Analysis of covariance is not mentioned, and experimental design receives only a half page. All this may be accepted by students and teachers in the social sciences, though its adequacy for their needs is surely doubtful. However this may be, it certainly will be unacceptable to most, perhaps all, other beginning students of statistics, though doubtless many will realize this only when they have ceased to be beginners.

---

* This review should not be construed as representing the official position of the U. S. Atomic Energy Commission.

The lack of systematic treatment of the analysis of variance and covariance is the more difficult to understand because it is so widely regarded not only as fundamentally important to applied statistics but also as so helpful to the author's stated primary objective of emphasizing the logic and understanding of statistics. Likewise, systematic treatment of experimental design would have been helpful to the same end, revealing as it does an entirely different aspect of the statistical whole, even if it were kept to only a few pages. Nevertheless, the author pursues his objective with some success, chiefly in the early chapters. He has done well, for example, in formulating some of the quasi-axioms of statistical thought, which can be so helpful to the tyro who remembers them, such as:

> The reader should develop the habit of appreciating that *an answer in terms of probability exists for every observed difference between two statistical measures,* even though he does not have the necessary ability, or perhaps data, to arrive at a precise statement of it. Just as the error in statistical statements is ubiquitous, so should be the awareness by the student of this fact (p. 78).

On the other hand, it should not be supposed that misleading and erroneous statements have been entirely avoided, as the following examples show. It is asserted (p. 73, using hypothetical temperature forecasts!) that the error is independent of the prediction "as it universally must be." There may be a definition of "error" which makes this statement true, but it need not be and usually will not be for real data analogous to the example. After a test for normality resulting in rejection of the hypothesis, we read, "knowing that successive daily maximum temperatures at a single spot are correlated and not independent, we have a ready explanation in that our items are not independent random samplings" (pp. 304–5), but of course this has nothing to do with the normality of distribution of either the sample or the parent population. A graph of the chi-square distribution for one degree of freedom (p. 285) shows an inflection where chi-square is about 0.13. Of t it is said that "It is the distribution of [means, differences of means, and of regression coefficients] . . . computed from small samples drawn from a parent normal population" (p. 286), obscuring the nature of this revolutionary statistic.

We are told (p. 226) that the expected value of the standard deviation is the square root of the expected value of the variance. There is the old error of "explaining" kurtosis simply in terms of the relative frequencies near the mean and in the tails (p. 242).

In addition to faulty statements, there are errors in formulae and tables, many of which are probably typographical, though none the less consequential to the student. For example, formulae for Fisher's fourth k statistic appear on pages 215 and 247, neither being correct, and a few errors were noted in the table (pp. 640–52) of normal probability functions.

The list of tables of statistical functions (including the usual mathematical tables) is excellent, each table being described as to content. It was a pleasure to the reviewer to find himself correctly credited for the table of the 20 per cent level of the z and variance ratio distributions. However, this pleasure was immediately mixed by inclusion of the 0.1 per cent table which of course is due to C. G. Colcord and L. S. Deming. The list of references, to both books and papers, would be much increased in value by elimination of such entries as Peters and Van Voorhis, *Statistical Procedures and Their Mathematical Bases* or, better still, by furnishing critical comments for each entry.

The notation seems unnecessarily cumbersome, and the terminology is sometimes unfamiliar, not to say strange. For example, there is frequent use of a symbol with both a bar and a tilde, the value found in a table is called the "consequent," as the value used to enter the table is called an "argument," and the reviewer is still wondering what the "xlnt" correlation chart is. This may be the appropriate place to remark that both "random" and "bias" are often used but never defined.

The author indulges in occasional flights of fancy as to subject matter or phrasing which serve to enliven the text, as when, in discussing significant figures, he considers a palsied man using a meter stick to measure a live angleworm.

The typography leaves much to be desired, both in appearance and accuracy, the spacing of lines, words, and letters being irregular, and typographical errors being frequent, especially in names. The index suffers from numerous errors of omission and commission, and additional cross-referencing would be helpful.

Mathematics of Statistics, Part One, Second Edition. *John F. Kenney* (Assistant Professor of Mathematics, University of Wisconsin at Milwaukee). New York 3: D. Van Nostrand Company, Inc. (250 Fourth Avenue), 1947. Pp. xii, 260. $3.75. (London W.C. 2: Macmillan & Co. Ltd. [10 St. Martin's St., Leicester Square]. 21s.) *Two reviews follow:*

REVIEW BY HOWARD LEVENE

Lecturer in Mathematical Statistics
Columbia University, New York, N. Y.

THIS is a slight revision, with minor clarifications and additions, of the first (1939) edition. The book deals almost exclusively with descriptive statistics, and requires no mathematics beyond high school algebra.

The first hundred pages deal with the mechanics and graphical representation of frequency distributions, the usual measures of central tendency and dispersion, and higher moments. These are followed by chapters on the normal curve, curve fitting, and correlation. The use of tables of the normal curve for curve fitting and graduation by areas is discussed. There are sections on fitting straight lines and second degree parabolas. An exponential trend is fitted by using the method of least squares on the logarithms, with a footnote warning that this is not the same as a least squares fit to the exponential itself. Two pages are devoted to fitting the Gompertz curve without an explanation of its use, while the logistic and Makeham's curve are merely mentioned and references given.

The summation notation is introduced early and used consistently throughout the text. This lends much greater clarity to the extensive discussion and proofs of various moment relations than is usual in a text on this level. For example, for k sets of variates combined into one set, the author obtains the relation

$$\sigma^2 = \frac{1}{N}\sum n_i\sigma_i{}^2 + \sigma^2{}_{\bar{x}_i}$$

where

$$\sigma^2 = \frac{1}{N}\sum_1^N (x_\alpha - \bar{x})^2,$$

and states that this will prove important in the analysis of variance.

This book is clear and well written and should be an excellent text for a course dealing with the material it covers. However, the reviewer questions the purposes served by such a course. The word "probability" appears for the first time on page 114 and is occasionally used thereafter, but a discussion of the concept is deferred to Part II (not here reviewed), and the notion of sampling and sampling error appears only in vague form in some of the discussions. The same letter is used throughout for both parameter and statistic. Sampling of attributes is not mentioned, presumably since a coherent discussion of it would require probability theory. Because of these and other deficiencies the book is not suitable for a terminal course in statistics. Students intending to go on to a "practical" course in statistics do not require such a thorough algebraic foundation, while those intending to take further courses in mathematical statistics might do better to devote their time to pure mathematics, so as to be prepared for an introductory statistics course using the calculus and probability theory. With the difficulty of finding time in the curriculum for all the subjects a student should study, it is doubtful if time should be devoted to material which must be largely repeated, in somewhat different form, in later courses.

Both the virtues and the defects of the book are well shown in the chapter on correlation. There is a good discussion of the nature of normal correlation. The normal bivariate surface is treated briefly, but the major mathematical treatment in this chapter is in terms of an existing finite sample.

Conclusions are drawn from this about the predictive value and the amount of relationship implied by a given correlation coefficient. While there is verbal reference to some vague "large" set of values or population, it would seem that this treatment would be confusing. Blakeman's criterion of linearity of regression is given, with a warning that it is not reliable and a reference to Part II for the analysis of variance test. The example on page 220 for which Blakeman's criterion is said to break down is poorly chosen, since the analysis of variance test is not applicable. Blakeman's criterion is the only significance test mentioned in the present volume, there being no mention that a criterion for existence of regression would be desirable.

There are a few minor errors, misprints and passages to which exception could be taken, but there is only one that the reviewer feels he should mention. On pages 147 ff. it is recommended that when two variables have a linear relationship and both are subject to errors, the sum of squares of perpendicular deviations from the fitted line should be minimized. This method is not invariant under change of scale of measurement. The problem of the determination of structural relationships is an important one, but requires a much more complete specification of the underlying chance mechanism, and even then cannot always be solved.

REVIEW BY RONALD W. SHEPHARD
Assistant Professor of Mathematics and Economics
New York University, New York, N. Y.

GENERALLY speaking this revision is not a substantial alteration of the previous edition—an elementary textbook on descriptive statistics, much used and a pedagogically sound treatment. The topics covered, their order and method of presentation, and the illustrative material are identical. The altered format of the revision results in a change of page numbers for corresponding textual material, which is a nuisance for those wishing to use it along with the earlier edition, but it has produced a better integration of figures and tables with related discussion. Still, the paragraph numbers are identical and one may avoid the changed page numbering.

It is difficult to find any changes in the textual discussion, and most students would use the revision along with the earlier edition without recognizing any differences. A new paragraph on variables and constants is added at the beginning of Chapter 1, and a not very useful paragraph on probability in Chapter 6. In the chapter on curve fitting a new paragraph is added dealing with logarithmic coordinate paper, and immediately preceding this Kenney has introduced a useful discussion of graphical determination of an exponential function with two new illustrative figures which change the ensuing numbering. Otherwise the alterations of test consist in a sentence here and there. The tables in the revision and earlier edition are the same.

One would naturally expect substantial change and improvement of student exercises, but this has not been provided. The alterations consist only in the addition of a few new exercises and constitute a disturbing deficiency of the revision for those looking for a new edition of this book.

More fundamentally, this revision is an example of a practice of publishers to rush into new editions at this time without having ready a thoroughgoing rewrite of their popular textbooks, resulting in little or no incorporation of new treatments and new applications of methods which have occurred during the war years. This is not to be taken as a criticism of Mr. Kenney's general method of treatment, which the reviewer has found to be a relief from the usual elementary textbook on descriptive statistics in the social sciences.

Descriptive and Sampling Statistics. *John Gray Peatman* (Associate Dean and Associate Professor of Psychology, City College of New York, New York, N. Y.) New York 16: Harper & Brothers (49 East 33rd St.), 1947. Pp. xviii, 577. $4.50.

REVIEW BY LEON FESTINGER
Research Center for Group Dynamics
Massachusetts Institute of Technology, Cambridge, Massachusetts

IN THE past two years there has been somewhat of a spurt in the publication of books on statistics, the reason for which is not too evident. Certainly many of the books on statistics which were published before the war had their little inadequacies, but the books published since the war seem on the whole to possess the same inadequacies or else to substitute deficiencies of their own for those of their predecessors.

The motivation for the publication of these new books does not seem to be the presentation of new material either. While the war period did produce a number of new developments in statistical methods, these are rarely included in the new books which appear. Perhaps the explanation lies in the differences of opinion, which no doubt exist, as to what to emphasize and how to teach statistics. The implication is certainly present, however, that a new book on statistics should be evaluated not only in terms of what it includes and how it presents its material but also in terms of the possible justification for using this new book rather than one which already happens to repose on the library shelf.

The title of the book under consideration, *Descriptive and Sampling Statistics*, correctly implies that the book treats separately the use of statistics for purely descriptive purposes—to summarize data—and the use of statistics as a tool for testing hypotheses concerning population parameters. These two sections of the book are of very different quality and must be evaluated separately.

The first part of the book, entitled "Descriptive Statistics" is, with respect to the content it covers, rather comprehensive. There is the customary array of frequency distributions, graphical representations, percentiles, means, medians, measures of deviation, and the like. There are some differences, however, between this and other similar texts. For one thing the discrepancies in terminology among different writers on statistics have led the present author to seek to change this terminology at various places. We thus find terms like "non-variable data" defined to mean what is usually meant by discrete data which does not necessarily form a scale, and we find the invention of new words such as vigintile, tercile deviation, and quintiserial correlation. Another difference which we find in this book is the somewhat heavier emphasis on the use of deciles and percentiles (which the author prefers to call centiles) than one usually finds. With this greater emphasis the author is able to point out quite clearly how the median is simply a particularly important centile value, and how the quartile deviation is one of a class of such deviational measures which might be computed.

In the preface to the book the author explains that he has "attempted a balanced presentation that will teach the student not only how to compute a statistical measure but when to use a particular technique and how to interpret a result." The degree to which this balance has been achieved, however, is open to some question. There is no doubt that a student who has carefully digested the material in the book will be able to compute the various measures which are presented. There are detailed steps for computation and numbers of illustrations of such computation. There is some doubt, however, whether this student will have an adequate enough understanding of what he is doing and why to insure the use of the proper measure where alternatives are available, and to insure adequate interpretation. In most places the book goes quite thoroughly into what to do but skims very lightly over the "why." For example, the reader of the book is told to correct a phi coefficient between two dichotomized variables by dividing by .637, but nowhere is there any mention of the reason for such a correction, the assumptions underlying this correction, or the conditions under which the correction should or should not be used.

The second section of the book, entitled "Sampling and Analytical Statistics," is of quite different character from similar sections in other textbooks. Some of these differences are commendable while others are rather unfortunate.

The present book is one of the very few which have attempted to include a discussion of sampling problems in public opinion research. The discussion is relatively brief, occupying only fourteen pages, but nevertheless does present the major techniques such as stratification, area sampling, and quota sampling. There is perhaps not sufficient critical evaluation of these techniques and not sufficient discussion of the problems of each. There is also, unfortunately, no presentation of sampling error formulae for dealing with these various kinds of samples.

Another feature of the book is a short but relatively good discussion of discrete probability distributions. This material, generally omitted from elementary statistics texts, includes important and useful statistical tools. This section, in the book under consideration, covers the theorems dealing with the addition and multiplication of probabilities and discusses the binomial expansion in detail.

Unfortunately, the remainder of the second part of the book is of rather inferior quality. The presentation consists almost exclusively of tests which might be justifiable when the number of cases in the sample is sufficiently large so that anything and everything distributes itself normally. The only exception made is for correlation coefficients where Fisher's z transformation is given. The t test is dismissed in one page with the idea that it is not very useful in psychology and the social sciences; the F ratio for significance of difference between variances is never mentioned; the techniques of analysis of variance are entirely omitted and the discussion of the chi-square test is rather poor. It may be that this selection of material was determined by the idea that this was to be an elementary text, but the reviewer finds difficulty understanding why outdated and inferior procedures are more elementary than newer and more adequate procedures.

It would undoubtedly be extremely difficult to write a textbook on statistics which was suitable for use in an elementary course without having some inaccuracies in the presentation. In the present textbook, however, there seem to be many inaccuracies which could be easily removed. We will list two of these doubtful statements so as to give the reader an idea of their nature.

> Thus, for a series of samples of less than 25 or 30 cases each, the form of the sampling distribution of any statistic derived therefrom will skew more and more from the normal probability curve, the smaller the size of the sample (p. 324).
>
> . . . *chance errors* are errors that are just as likely to occur as not to occur; hence they are equiprobable (p. 332).

Whether or not readers will find *Descriptive and Sampling Statistics* a worthwhile text to use will undoubtedly depend largely upon their personal preferences as to what should be taught in an elementary statistics course and how it should be taught.

REVIEW BY FREDERICK F. STEPHAN

Director of Study of Education and Professor of Statistics
Princeton University, Princeton, N. J.

How should a textbook be appraised—in terms of its competition or in terms of the needs it should satisfy? Compared with other texts in current use, this book has several commendable features that will appeal to some teachers and students. For the most part, these features consist of interesting discussions of the meaning and applications of statistical methods in psychological testing and public opinion polls. Among them one may mention an interesting sketch of the historical background of statistics and several passages in the sections on sampling, test validity, and correlation that reveal considerable skill and ingenuity in presentation. However, there is another standard by which a textbook may be judged, namely, its effectiveness in remedying the defects of other books and bringing the teaching of the introductory course abreast of recent developments in basic theory and method. On this standard, the book does not appear to be satisfactory. To too great an extent it seems to be a rewriting of the content of familiar textbooks, principally Peters and Van Voorhis, and parts of Croxton and Cowden, and Yule and Kendall, rather than a fresh and modern formulation of the introductory course based on current research applications and methodological advances. It even falls behind these books (published some ten years earlier) in many respects. For example, the reviewer could not find as much as a casual reference to the analysis of variance.

Of course, there is plenty of room for disagreement about what is appropriate to a first course, even one of a year's duration, especially when it is addressed. as this is, to students with no college mathematics. Yet, if the author really expects to give students "the ability to select the most relevant statistical technique and to make the necessary computations with a minimum of error" (p. xvii), he could well include a discussion of such topics as the efficiency of estimating functions and the power functions of statistical tests, at least to the extent of providing a clear statement of the problem to be solved in each case. Certainly these concepts are fundamental to an understanding of the logic that underlies the application of statistical methods and they can be presented simply enough to students who can understand the great variety of correlation methods that are discussed in great detail.

While the book is intended for research workers in the social and biological sciences, it neglects important parts of these fields and emphasizes psychological and educational testing and public opinion polling. It slights psychological experimentation, genetics, and other major regions of the application of statistical method to research in these fields.

The book is divided almost equally between "descriptive" statistics and "sampling and analytical statistics." As a consequence the normal distribution and correlation are discussed extensively before sampling and probability. This separation is harmful, in the opinion of the reviewer, for it obscures the extent to which statistical theory has been unified.

The topics that are covered include: categorical data, proportions, ratios, index numbers, percentages, correlation of categorical data, frequency distributions, centiles, mean, standard deviation, the normal curve, correlation, sampling, probability, tests of significance, chi square, prediction from correlations, evaluation of psychological tests, and cluster and factor analysis.

The author took pains to avoid sacrificing accuracy in attempting to write clearly but there are a number of places at which statements are confused or incorrect when read literally. A few examples will be cited:

An IBM card is illustrated and described as having ten positions in each column (p. 38). No mention is made of the x and 12 or v positions.

"The arithmetic mean . . . and the standard deviation . . . The method of their computation is essentially algebraic and is often described as the method of moments" (p. 150).

"The normal distribution . . . is a unimodal distribution with the greatest number of frequencies at the mean" (p. 175). The discussion of the normal distribution is inadequate; ordinates are given as probabilities (p. 184).

If a test of significance does not indicate that the difference is significantly greater than zero, "we can be confident that no real differences exist in the characteristics compared" (p. 401).

There is a meager one-page bibliography. It does not include such books as Snedecor or Lindquist, which are needed to cover important fields that this book slights. The appearance of books like this emphasizes the importance of more effective communication between statisticians in various fields and improvement in the teaching of statistics, two functions in which the American Statistical Association should be increasingly active.

Fundamentals of Statistics. *J. B. Scarborough* (Professor of Mathematics United States Naval Academy, Annapolis, Md.), and *R. W. Wagner* (Assistant Professor of Mathematics, Oberlin College, Oberlin, Ohio). Boston 17, Mass.: Ginn and Co. (Statler Building), 1948. Pp. vii, 145. $2.50.

REVIEW BY LEO AROIAN
Assistant Professor of Mathematics, Hunter College
New York 21, New York

THE appearance of a new text on statistics is always a source of pleasant anticipation for the harassed teacher searching hopefully for the ideal text. The present text, intended for engineering and science students, found its birth in the need for a short course of twenty lessons at the United States Naval Academy. An elementary knowledge of calculus is assumed. The book is remarkably brief. Simplicity and clarity are stressed throughout. The eight chapter headings and appendices are representation of data, averages, measures of dispersion, comparison of distributions, correlation, probability functions, the normal curve and a generalization, sampling, one appendix on probability, a second one on the tables of the normal curve and the second, third, and fourth derivatives of it, a third one listing nine texts on statistics recommended for collateral reading, and five pages of answers to problems.

The originality in conception of the book must be highly praised. Undoubtedly it will be extensively imitated. The mathematics carries the burden of the explanations. Very few words are wasted. The authors exhibit considerable ingenuity in the development of the topics and are at their best in descriptive statistics, excellent in their treatment of probability and sampling once their point of view is granted. They have not been afraid to drop many topics considered sacrosanct such as Sheppard's corrections. Many problems illustrate the main points of the text. A nice balance is maintained between theory and numerical computation. Yet there is a greater nod to the past than to the future.

Fundamentally the reviewer disagrees with the authors' philosophy of testing hypotheses. Some intuitive explanation of the modern theories of probability as developed by von Mises and Kolmogoroff would be desirable. In many places too little explanation is given. Only two-sided tests of hypotheses are mentioned and needless to say nothing concerning Type II errors, the error of accepting a false hypothesis. For sample means, N large, this is a simple matter. More serious is the confusion of two entirely different problems: "Given a probability function, all parameters known including the population mean, what is the probability of obtaining a sample mean $\bar{x}$ greater than x_0?" with the problem "A sample of N variates is drawn, N large, from an unknown population (possibly of known functional form), a statistic is calculated say the sample mean, what can be said of the population mean?" On page 106 the statement is made: " $\cdots N_{pop} = N_{samp} \pm \frac{\sigma_{samp}}{\sqrt{n-1}}$. This is interpreted to mean that the probability is about 2/3 that the mean of the population is within the range $M_{samp} - \sigma_M$ to $M_{samp} + \sigma_M$. "(!) Actually after a sample is drawn the probability that the population mean should lie in such an interval is zero or one. What is needed here is the theory of confidence intervals, certainly a fundamental topic.

Some important topics omitted are significant figures, the range, *J*-shaped curves, the chi-square test of goodness of fit, analysis of variance, control charts, and small sampling theory. Topics which might readily have been omitted are the quartile deviation, the coefficient of curvilinear correlation, and the coefficient of rank correlation. On page 24, $\sigma_x = a\sigma_y$ should read $\sigma_x = |a|\sigma_y$. The common error concerning α_4, the measure of kurtosis, occurs on page 33. The attempt is made to justify their statements by use of a Gram-Charlier Type A of three terms including α_4. As Cramer remarks in his recent text (p. 184, p. 230) an additional term $\frac{10\alpha_3^2}{6!}\phi^{(6)}(x)$ is required; consequently the authors' remark is not accurate even for a Gram-Charlier series in which the terms are grouped according to the same order of magnitude. Naturally in some particular cases $\alpha_4 > 3$ is associated with a high peak at the mean. The final discussion on the limitations of the coefficient of correlation is good.

A word might be said about notation. The letter x is used to represent the original variable, a variable with mean zero, and a variable with mean zero and standard deviation 1. The mean is usually written as M whether referring to a sample mean or a population mean, although occasionally a distinction is made as on page 105. The same remark applies to the standard deviation and the measures of skewness and of kurtosis.

It is a pity that in the section on collateral reading no mention is made of the recent books by H. Cramer, H. A. Freeman, P. Hoel, A. Wald, C. E. Weatherburn, and S. S. Wilks.

Practical Rules for Graphic Presentation of Business Statistics. *L. Edwin Smart* (Professor of Economics) and *Sam Arnold* (Assistant Professor in Business Research). (Ohio State University.) Columbus, Ohio: Bureau of Business Research, College of Commerce and Administration, Ohio State University, 1947. Pp. xii, 86. $2.00.

REVIEW BY RUDOLF MODLEY
President, Executive Research, Inc.
350 Fifth Ave., New York 1, N. Y.

THE authors of this small book have set themselves the task to write a simple "How to Do" book on the graphic presentation of statistics for businessmen and students. In going about their job they have set forth specific rules applying to all charts and rules applying only to some types of charts. Each text discussion is accompanied by one or more illustrations. The book deals with bar charts, pie charts, line charts, three dimensional charts, pictorial charts, and maps. The last third of the book presents samples of business charts. In presenting text and charts the authors do not discuss why something should be done in a certain way. They rarely present alternative methods of presentation; they only give the rules and leave it at that.

There can be little doubt that a simple book of this nature is badly needed today. The use of charts has been increasing rapidly and many businessmen need quick and simple guidance on how to draw charts. Unfortunately, "Practical Rules" cannot be recommended as filling this need. There are several reasons why it does not.

a) The proof of a chart is in its looks. The least we can expect from a book on graphic presentation is that the charts are well laid out, well designed, and fit harmoniously into the book. This is not the case in the book under review. Nearly all the charts look awkward; they lack proper balance within themselves and on the page. There is poor selection of size for lettering and poor weight selection for rulings and curves. The instructions on lettering sizes (p. 6) are insufficient, instructions on grid and curve weights (p. 34) limited to "light" and "heavy." Unnecessary or incorrect rules, as for instance, "that the grid will be wider than it is long in somewhat the same proportion as 5 is to 3" (p. 7) or that pictorial symbols should be "mere outlines with little detail" (p. 48) also make it more difficult to obtain clear and forceful charts.

b) This leads us to the second shortcoming of the book. Instructions are usually given as if there were "generally accepted standards" of chartmaking. While there are some such standards it is obvious to all of us who deal with charts that effective presentation calls for considerable flexibility. Most of us will make trial layouts before we draw up a chart. We do so because this permits us to make adjustments in the many different elements which go into a chart. We reorganize or eliminate some elements, we emphasize others, we adjust dimensions and sizes of lettering until we are satisfied that the chart is truly effective. Only then do we begin the final draft. If the authors had tried harder to make the draftsman understand what one can "get out" of a chart by different treatments instead of giving a set of seemingly inflexible rules, a more helpful book would have resulted.

c) The authors state that they did not want to "give the pros and cons of various practices." (Actually they have done so occasionally displaying a preference for 100 per cent component bar charts over pie charts and absolute component bar charts.) It seems to this reviewer that an objective listing of the advantages and disadvantages of certain types would be essential in a practical book for those who cannot be expected to judge all the elements which go into making a chart correct and effective. Alternative ways of presenting the same set of data would also have been helpful. The wealth of samples included in Brinton's *Graphic Presentation* is a considerable help to the layman. It gives him a chance to select from different methods the one which fits his needs most clearly.

d) This reviewer has spent about twenty-five years in pioneering the development of pictorial presentation of statistical data. The shortcomings of the book under review naturally become most clearly obvious to him at this point. Pictographic presentation is advantageous in a limited field. These limitations in regard to audience, accuracy, and cost should be stated if we are to give a prospective chartmaker practical help. If the chartmaker decides for pictographs he will have to follow a set of simple rules; the authors give one of these and leave out all others. Several ways are available to duplicate pictorial symbols so that charts can easily be prepared. They can be photostated, photo-offset, type-stamped, or rubber stamped. Symbol sheets are sold by some organizations. None of this is mentioned in the book. The stencil suggested in the book is not being used by any chartmaker known to this reviewer. The statement that "pictorial charts are more effective if the classification is qualitative" is not correct. There is ample literature in existence on pictorial presentation to have avoided the textual and graphic shortcomings in this section.

e) A short book of this type needs a bibliography. Many readers will want to know where to turn for more information.

In 1926 the American Society of Mechanical Engineers formed a Sectional Committee on Standards for Graphic Presentation. In April 1936, the Society published a *Code of Preferred Practice for Graphic Presentation—Time-Series Charts.* On the basis of comments and suggestions received, a revision, *Time Series Charts—A Manual of Design and Construction,* was published in November 1938 with the approval of the American Standards Association. Nevertheless, the manual is not a compilation of rigid rules. The Committee believed that effective presentation calls for flexibility of treatment rather than for standardization, that each chart should be individually planned with due reference to the special characteristics of the data and the particular use to which the chart is put. Within these limitations there is now available a set of well-tested rules for the presentation of time-series charts. The authors could have performed a real service to students and businessmen if they had tried to popularize the more important elements contained in the manual. At present this little volume, *Time-Series Charts*, is still the best guide for those who want to draw charts of the conventional type.

Control Charts: An Introduction to Statistical Quality Control. *Edward S. Smith* (Professor of Mathematics, University of Cincinnati, Cincinnati, Ohio), New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1947. Pp. xi, 161. $3.00. (London, W.C. 2: McGraw-Hill Publishing Co. Ltd. [Aldwych House, Aldwych]. 15*s*.) *Two reviews follow:*

REVIEW BY JOSEPH M. JURAN

Professor of Administrative Engineering and Chairman of the Department
New York University, New York, N. Y.

HERE is another of the numerous books now emerging to expound the use of statistical methods to control quality of product.

There is a good discussion of the use of control charts for averages, ranges, and number or fraction of defectives. Visual aids in the form of charts or tables are well prepared. The exercises are well chosen.

The claim that the book is "a test for the busy executive" is wishful thinking. In this work, and for that matter, in others, there is little evidence that the authors have truly grasped the quality problems of the industrial executive.

Industry's problem has been and still is to make product to meet specifications. The statistical methods, if they are to be grasped by industry, must be presented as a means to the solution of that problem. The emphasis must, in the first instance, be *not* on (*a*) how to achieve controlled processes; the emphasis must rather be on (*b*) the fact that through a controlled process it is easier to make products to meet specifications.

To a statistician, (*b*) seems only too obvious; ergo the executive should take (*b*) for granted, and thus only (*a*) remains to be learned. The precise opposite is true. The practical problem is to convince the executive of the truth of (*b*). Even after he is convinced of (*b*) he will ordinarily not be concerned with (*a*); instead he will delegate the learning of (*a*) to another.

With the exception of the unrealistic first chapter Professor Smith's book can well serve as a supplemental text for undergraduate engineers.

REVIEW BY MASON E. WESCOTT

Assistant Professor of Mathematics, Northwestern University
Evanston, Illinois

PROFESSOR SMITH has long been a successful teacher of college mathematics. During and since the war, he has participated as an instructor in numerous 8- and 10-day intensive training courses in statistical quality control for industrial people. He has conducted several successful in-plant training courses and has served as technical advisor for the installation and operation of statistical quality control programs. Drawing on this background of theoretical and practical experience, he has contributed *Control Charts* to the McGraw-Hill Industrial Organization and Management Series.

It has been the author's intention to write a text "for men and women who wish to acquire knowledge of the methods and charts of statistical quality control, but who are too busy to have an interest in the mathematical aspects of the methods." He believes that "it is more important than ever that everyone concerned with such production (production of goods of high quality) should be able to make and read control charts."

With the latter point of view, this reviewer is in complete agreement, but he does not feel that such understanding can be acquired without more exposure to things mathematical than the "busy executive" generally feels he can or wants to endure. As a matter of fact, the innocent reader who begins Professor Smith's book expectantly hoping to be painlessly initiated into the mysteries of statistical quality control, soon finds himself face to face with formulas and exposition that are pretty likely to appear frighteningly mathematical to him. That this is so, should not be held against the author, because Professor Smith has very nearly succeeded in keeping to a minimum the formal aspects of his subject. Rather, it is convincing testimony to the fallacy of trying to undertake an expository discussion of a technical subject, be it even so simple as control charts, without sooner or later having to use symbols and language that to a great many nonprofessional people must seem decidedly mathematical.

Chapter 1 is probably the best chapter in the book from the point of view of the audience to whom the book is addressed. Here the author introduces the definitions he proposes to follow for such concepts as "control," "quality," and "satisfactory control of quality," distinguishes well between inspection for the purpose of *controlling* quality and inspection which merely sorts good product from bad, outlines pointedly the fundamental role of a control chart as a criterion for telling when to take action and when to leave well enough alone, and lists ten benefits to be expected from a properly administered control chart operation. One may very well take exception to parts of the definitions given, but on the whole, they are acceptable for the purposes intended. The statement that "*A production process is said to be producing goods of satisfactory quality if 99.7 per cent of the items are satisfactory in quality*" is, to this reviewer, the kind of a statement he would rather not see used with the emphasis it is made to carry here.

The remaining chapters are, for the most part, well organized and clearly presented. One becomes conscious of a tendency to emphasize details rather than to paint the picture in broad outline. The latter treatment would have seemed more consistent with the objective of "selling" the subject to an audience which is not supposed to be interested in details. However, Professor Smith is understandably more the teacher than the salesman in these chapters, and as such, he anticipates and tries to remove the mental hazards that experience has taught him block the student's progress in learning how to make and use control charts. Certainly, if the reader will exercise the patience and intellectual curiosity to study through these chapters, working the problems proposed as he goes, he will end up with at least the mechanical skill to make a control chart, and a somewhat better than bowing acquaintance with the operational meaning of such a chart.

In these chapters a number of technical points appeared questionable to this reviewer, but most of them are not serious enough to warrant detailed discussion here. Three points, however, cannot be ignored.

a) There is an irritating looseness with which the word "normal" is used. Such phrases as "normal, symmetrical distribution," "normal average," "normal and approximately symmetrical frequency distribution," "normal process" occur repeatedly without the benefit of any consistently adhered to definition; in fact, without any "definition" at all! Unless the reader has unusual insight or sophistication, he may very well end up with the happy delusion that everything will be just dandy if only it can be described as "normal." The word is simply overworked in the last four chapters without reserving to it the distinction it rightfully possesses in statistics.

b) It is the opinion of this reviewer that on pages 48–52, and elsewhere in subsequent discussions, Professor Smith overplays his hand in the matter of testing control against product specifications. In the first place, the practice of putting "danger limits" (three-sigma limits for individuals) on a chart for averages is questionable, particularly if the chart is to be kept at the machine. Exhibiting the to-be-expected spread for *individuals* on a chart for *averages* simply invites trouble, because the operator finds it hard to understand why he must keep the points within the narrower band set by control limits for averages when the "danger limits" are way out beyond this band.

In the second place, the discussion manages to convey the impression (although the contrary is explicitly stated on page 52) that there will frequently be hair-splitting closeness between a "danger limit" and a blueprint specification which has to be resolved before a controlled process can be adjudged in satisfactory control. The criterion described is not in error: it simply overemphasizes a relatively unimportant point. Time might better have been spent in showing how to estimate the per cent of production likely to be out of tolerance when control is not satisfactory, and in discussing more fully the problem of balancing the cost of sorting and/or rework against that of making process or specification changes that will yield control safely within acceptable blueprint tolerances once the optimum capabilities of the process have been determined.

c) The discussion in Chapter 4 uses the word "lot" where the word "sample" would ordinarily be preferable. The beginner reading this chapter must surely experience a sense of confusion when, after digesting 21 pages in which he has had the word "lot" drilled into him at every stage, he suddenly discovers on page 106 that a *box* full of beads is a "lot," but a *paddle* full of beads is a "sample," and that now it is the *sample* that is to be inspected 100 per cent rather than the *lot* as he has previously been led to expect! Evidently the author intended this chapter to deal with attribute inspection for process control rather than for acceptance sampling, although, of course, the two are not mutually exclusive. But if inspection for process control is intended, it would seem more consistent and realistic to characterize a "day's production" (which the author says must be inspected 100 per cent!) as a "sample" from the cause system represented by the production stream than it is to describe these items as a "lot."

Furthermore, it would be entirely possible for the beginner to get the impression from this chapter that 100 per cent inspection (in the ordinary sense) is required to keep a fraction defective chart. Many plants exercise very satisfactory process control through the maintenance of fraction defective charts based on samples that are a very long way from being what the average man has in mind when he finds himself reading about 100 per cent inspection of a *lot*. It is the *sample* that is inspected 100 per cent, and it would be wiser to say so, rather than to use language that can be so easily misinterpreted.

Table II appears to be potentially useful in simplifying the calculations required for control limits on p and np charts. Its use is amply illustrated in Chapter 4.

The typography is good; tables, charts, and diagrams are well done and legible. The book is easy on the eyes. Except for a few missing punctuation marks, only two errors were found: on page vi, the name should be Paul Peach, and on page 142, the bar has been omitted over the R in the third line from the top.

Control Charts could prove useful as a text for a short training course at an elementary level. As a textbook for untutored home-study or as a "quickie" for the casual reader who wants to see the forest but not the trees, it is the opinion of this reviewer that the book is likely to prove disappointing.

Méthods Statistiques Modernes des Administrations Fédérales aux États-Unis. *Pierre Thionet.* Paris: Hermann & Co. (6 rue de la Sarbonne), 1946, Pp. 117.

REVIEW BY W. EDWARDS DEMING
Adviser in Sampling, Bureau of the Budget
Washington

THIS book is a unique and definite contribution toward improved statistical practice, as well as a delight to read. The author exhibits unusual ability in theory and extraordinary clarity in presentation. He recognizes the need for proper field organization in conducting sample surveys, and appreciates the need for reliable figures in government administration. The study of this book has been a pleasing experience for the reviewer. It is interesting to see how we are viewed from across the Atlantic, and gratifying to find the essential points of progress in the United States so clearly pointed out. Here, we are too close to give a coherent account of what has been happening. Moreover the reviewer feels a personal debt to Mr. Thionet for providing in one cover a clear exposition of much published research that was difficult in its original form.

The book was perhaps intended mainly as a digest of the essential mathematical papers that have made impact on statistical practice in the United States. Actually, it is much more than that. It is a manual as valuable to the nonmathematical administrator as to the student of theory. The author pleads for the use of figures in government administration in place of abstract theorizing and political maneuvering. Now that sampling has been developed to its present stage of reliability, real figures instead of fiction are possible and there is little excuse for operating without them. Progress is not achieved by shackling the theoretical statistician, but by giving him a chance. The author is fearless: "Et cette évolution de la méthologie statistique et même de la statistique mathématique risque de rester mal connue en France où les personnes (déjà peu nombreuses) qui connaissent et utilisent la méthode statistique, ne s'intéressent trop souvent qu'à l'un de ses aspects: pratique, théorique, mathématique, administratif, etc., alors que nous entrons dans un domaine où ils s'interpénètrent tous . . . En France, où d'ailleurs le nom même de statisticien, jadis équivalent de celui d'actuaire, a disparu de la hiérarchie administrative officielle, les conditions sont profondément differentes . . . Et l'ensemble de la nation fait preuve d'une prédilection marquée en faveur des idées abstraites de préférence aux réalitiés chiffrées, de sorte qu'un gouvernement ou une assemblée se préoccupera rarement de l'avis des statisticiens quand une décision doit être prise" (pp. 4–5).

This book is destined to be widely used. Fortunately a large part of the world reads French. With the coming census of population and agriculture, which most countries have agreed to take in 1950–51, there is an appalling need for sampling. The most serious handicap is official deafness, but lack of trained personnel and suitable texts is almost as serious. No one book will meet all needs, but every good one will help. In many countries, proper sampling procedure offers the only hope of obtaining census information of sufficient reliability and speed to be useful. Even in this country, with a well-developed census, sampling is used not only for current information on employment, housing, cost of living, characteristics of farms and farmers, and many other purposes, but was also used concurrently with the complete censuses of 1940 and 1945 (and will likewise be so used in 1950) with the aim of broadening the scope of the census and speeding up the returns.

The student of theory will find in this book an ably written and unified account of the mathematical researches that have guided progress in statistical practice. It is impossible to portray either its brilliant selection of contents or its lucidity in a review, but a hasty summary would include a nonmathematical introduction to the general principles of sampling, its advantages and disadvantages; descriptions of the U. S. Census tabulations and their use in sampling; appreciation of the fact that sampling is not only an expedient of war-times but an instrument of stability in peace; some general theory and remarks on application. Then follow five chapters and two annexes, in which are beautifully expounded in uniform notation a number of methodological researches that have had impact on modern practice. Works of Hansen and Hurwitz, Neyman, the Madows, Cochran, Jessen, Stephan, Stock, Frankel, Geary, Hauser, Snedecor, King, H. Fairfield Smith, and Walter Hendricks are included. Unfortunately the author missed the profound researches of Mahalanobis, whose reports on sampling for the acreage of jute and other crops and conditions in Bengal have been recorded from time to time since 1937 if not earlier. Mahalanobis's magnificent article "On large-scale sample-surveys" (*Philos. Trans. Royal Soc.* 1944) probably reached Paris a bit late for inclusion: this article not only described Mahalanobis's brilliant researches in the use of grids in agricultural sampling, but also lays down the framework into which all sampling procedure must hereafter be fitted. This framework consists of four steps:

> *i*. Decide in advance on the widest permissible sampling tolerance, and specify also what measurements are desired on the differences between various alternative questionnaires, definitions, methods of training, and between interviewers, respondents, etc.
>
> *ii*. Design the cheapest plan that will meet these requirements and stay within the limitations of personnel and physical equipment available and any administrative restrictions that must be imposed. Compute the expected cost.
>
> *iii*. Reconsider Steps *i* and *ii* in view of the expected costs and precision. Perhaps relax the prescribed precision or forego some of the aims to cut costs; or tighten up the precision if the cost is not too great.
>
> *iv*. After the results of the survey are in, compute the costs and the sampling errors and the differences between the various alternative procedures that were specified. This appraisal of the results is independent of the assumptions that were made in the planning. It provides further knowledge of the proportions and variances in the universe, and of costs, so that a cheaper job can be done next time.

It is to be noted that these steps, which constitute modern statistical practice, were impossible and unthinkable until the theory of probability was introduced into the design of surveys and experiments. Thionet is clear on these points, although the steps as enumerated above are not stated explicitly.

The author himself has made a number of notable contributions in theory, not only by filling in gaps but by extension as well. By coincidence these extensions sometimes follow lines of unpublished recent researches that are even now profoundly modifying federal statistical practice. One such contribution is in the topic of *estimation*, in which the author does not shrink from the use of biased estimates. Given the simple case of Bernoulli sampling from each of L urns, three plans of estimating p are discussed from the standpoint of consistency, bias, and variance (pp. 28–30). If x_i is the number of black balls in a sample of n_i from the ith urn, three possible estimates are these:

1. $\hat{p} = \dfrac{\sum x_i}{\sum n_i}$
   - a. Consistent
   - b. Biased unless n_i is proportional to N_i
   - c. $\text{Var } \hat{p} = \dfrac{1}{n^2} \sum \dfrac{N_i - n_i}{N_i - 1} n_i p_i q_i$

2. $\hat{p} = \dfrac{1}{L} \sum \dfrac{x_i}{n_i}$
   - a. Inconsistent
   - b. Biased
   - c. $\text{Var } \hat{p} = \dfrac{1}{L^2} \sum \dfrac{N_i - n_i}{N_i - 1} \dfrac{p_i q_i}{n_i}$

3. $\hat{p} = \dfrac{\sum N_i \dfrac{x_i}{n_i}}{\sum N_i}$
   - a. Consistent
   - b. Unbiased
   - c. $\text{Var } \hat{p} = \dfrac{1}{N^2} \sum N_i^2 \dfrac{N_i - n_i}{N_i - 1} \dfrac{p_i q_i}{n_i}$

Then follows a discussion regarding comparisons of the three variances, with and without stratification, and under various systems of apportioning the sample amongst the strata. The author makes it clear that the aim in making an estimate is to come forth with a number whose reliability can be controlled and measured, whether it be biased or unbiased; that the procedure of estimation is part of the sampling plan; and that the cost of the sample and the precision of the results will depend on the procedure of estimation as well as on the method of selecting the households that are to be interviewed.

A review is not complete without one criticism. The reviewer fears that the section-heading on page 13 "Pas d'échantillonnage sans l'aide des recensements" will lead the uninitiated to suppose (as so many of the laity now suppose) that a sample must rely on a complete census for the total number of people; that the sample can only give estimates of the proportions of people in various classes, and that, even then, it is well to have some census information as a check. Information from a previous census is indeed useful in designing a sample, as Thionet points out, but it is not necessary. Clever use of available census information for stratification, for ratio-estimates, and for drawing areas with probability in proportion to previous size, will be effective in decreasing the standard error of subsequent samples, and hence in cutting costs. But census information is not necessary; detailed maps can be used for defining areal units and are often so used even when census information is available. Obsolete census information and obsolete directories, civil and parish lists, are useful for sampling because, with proper techniques of estimation, the errors in the prior information cancel out, leaving the sample estimates unbiased.

Mr. Thionet deserves the praise and thanks of all statisticians. The reviewer closes with the hope that translations of the book into Spanish and Chinese will appear in time for assistance in planning the world census of population and agriculture in 1950–51 in places where French is difficult to read. An edition in English would likely find high favor as a text and reference book not only here but in other English-speaking countries, notably in countries like India where statistical teaching and practice have reached high levels.

An Extension Course in Statistical Quality Control. *A. E. R. Westman* (Director, Department of Chemistry, Ontario Research Foundation, Toronto 5, Canada). Reprint of series of articles which appeared in *Manufacturing and Industrial Engineering*. Toronto 1, Canada: Consolidated Press Ltd., 1947. Pp. 58. Paper. $1.25. *Two reviews follow:*

REVIEW BY A. C. COHEN, JR.
Associate Professor of Mathematics
University of Georgia
Athens, Georgia

THIS syllabus was originally prepared for use in statistical quality control courses which Dr. Westman conducted during the war under sponsorship of the Canadian Government for the benefit of Canadian industry. The Canadian courses were closely patterned after those offered in this country by OPRD for American war industry, but as reflected by the syllabus, the Canadian courses were even more condensed than the intensive OPRD courses.

Frequency distributions, control charts for measurements, control charts for counts, acceptance sampling with particular emphasis on the use of the Dodge-Romig tables, installation of a statistical quality control system, and other related topics are among the subjects discussed in the thirteen two-hour lessons which comprise the course. The various lessons include lectures to be presented by the instructor, demonstrations in which both the instructor and class members participate, and laboratory periods during which class members execute assigned exercises under the instructor's supervision. The numerous demonstrations which are described in detail effectively illustrate the theory of probability as it concerns statistical quality control, and the many exercises afford the student practice in applying the statistical techniques of quality control to practical situations.

This publication will unquestionably be a welcome addition to the quality control library of anyone concerned with the training of quality control personnel. It is particularly recommended to anyone confronted with the task of setting up and conducting a course of instruction in statistical quality control at the elementary level. Unfortunately its usefulness is marred by a somewhat lengthy list of errata. Most of the errata are typographical, but several involve formulas which if not corrected will certainly prove confusing and annoying to the uninitiated reader. A more clearly stated distinction between sampling for process control and sampling for acceptance (i.e. disposition) of product would be desirable. In this connection, although Lesson 12 is concerned with acceptance sampling, that fact is not explicitly stated, and the reader having no previous knowledge of the subject will likely find the omission confusing. The subjects of correlation and contingency tables which are included in Lesson 13 may prove too advanced for a first course in quality control for anyone who has not had previous training in statistics, but if desirable, these topics can be omitted without detriment to other material in the course.

REVIEW BY PAUL PEACH
Associate Professor, Institute of Statistics
The University of North Carolina, Raleigh, N. C.

DURING the war, in connection with the 8-day courses in statistical quality control organized by the U. S. Office of Education and the War Production Board, there was prepared a study manual by Professors Holbrook Working and Edwin G. Olds which more or less stabilized the pattern of instruction for those courses. That manual was never published, but several of its legitimate children have attained the dignity of print. Dr. Westman's pamphlet contains an introduction in which this parenthood is specifically acknowledged.

The student begins by making some histograms and calculating a few means and standard deviations. Three following lectures deal with control charts for measurements, three more with the charts for p and c, and then three with acceptance sampling. The last three bear the titles "Installing a Quality Control System," "Other Sampling Inspection Schemes," and "Introduction to Advanced Topics."

The presentation is, I think, an improvement over previous ones; the style is clear, and each lecture is illustrated by diagrams, tables of data, and similar aids. A laboratory or demonstration period is provided for at the close of each meeting. Most of the chapters conclude with a list of special terms, though without definitions. Taking the work as a whole, one might say that for the most part it follows the pattern which was created in the years 1942 to 1944, treating the customary subjects in the established way, with just about enough novelty to justify a separate existence.

There may be some difference of opinion about these novelties. Dr. Westman includes less theoretical statistics than did the old WPB course, and I should be inclined to go even farther and eliminate theory altogether. The avowed viewpoint of the pamphlet is that statistical quality control is a branch of industrial engineering, and while I disagree with this idea, yet I agree entirely that a short course for engineers should skip the statistical theory. A noteworthy recent contribution to the literature is the new British Standards Institution publication *Fraction-Defective Charts for Quality Control* (B.S. 1313:1947) which has been almost completely purged of statistics.

Dr. Westman's final chapter treats control charts for composites, correlation, and the chi-square test. Presumably this chapter is intended as a sort of bonus, since it seems unlikely that any lecturer would attempt to cover this much ground in a single period. Here again we may ask whether, when we can at best do so little, it would not be better to do nothing. Nevertheless, since the chapter comes at the end, the instructor can if he likes omit it, and no harm is done.

This last chapter means, I suppose, that engineers realize that the usual introduction to control charts and acceptance schemes is only a beginning, and that the value of other statistical tools is becoming recognized. Unfortunately, most of the present literature on statistical quality control ends in a blind alley; the student who wishes to learn advanced methods has to retrace his steps and start all over with new definitions, new concepts, and a new notation. In his first twelve chapters Dr. Westman follows the current writings on control charts, and of course is in a difficulty in his thirteenth. He cannot escape introducing the "degrees of freedom" concept, and gives formulas for the number of degrees of freedom appropriate to correlation coefficients and contingency tables; I am afraid that with this much introduction the concept will still seem pretty esoteric to his readers. There is room in our educational system for a manual that will do what Dr. Westman has done, but with a viewpoint and notation that will enable the student to study further if he likes without discarding most of his elementary instruction.

A word about typography may be in order. Most printed works, especially first printings, contain typographical errors; with these I am not concerned. Print shops not accustomed to mathematical text usually have trouble with it, especially with equations and other algebraic expressions which must be set by hand. Frequently, if special mathematical types are not available, it is necessary to use makeshifts. A lot can be done in this way if the author and the typesetter work closely together; but in the present work the handset portions are really a mess. Dr. Westman was also done dirt by in other ways; for instance, his publishers have not even bothered to use uniform style for chapter headings. The presswork and paper are good, but some of the editorial work is plain slovenly.

Toleranser och Passninger. *Stig Zetterlund.* Stockholm, Sweden: Sveriges Mekaniska Verkstaders Forbund, 1945.

REVIEW BY CECIL C. CRAIG
Professor of Mathematics, University of Michigan
Ann Arbor, Michigan

OF THIS rather compendious book on tolerances and fits written by a member of the staff of the famous Swedish manufacturer of ball bearings, the SKF company, only chapter 8 is of general interest to a statistical audience. To those who have devoted at least a share of their energy in recent years to advocating the use of statistical methods in the control of the quality of articles produced by repetitive processes, it is gratifying to see this much recognition given in a book like this to the importance of probability and frequency laws, but it is also disappointing to see that the development here still stops well short of adequately conveying to those concerned with manufacturing interchangeable parts how useful statistical tools can be to them.

In this chapter the fact is brought out that the actual sizes of a part produced to the same nominal dimension are subject to variations which, if due to chance only, tend to be distributed in a bell-shaped frequency curve, very often usefully summarized by the normal frequency law, hence by means of the average and standard deviation. It is well illustrated that the play between two assembled parts, such as a hole and an axle, when each is subject to random variation, even of a rather arbitrary sort, tends to be so distributed. The chief problem actually dealt with is that of the variability to be expected in the sum of the dimensions of two assembled parts, when the variability of each is known.

Conference Papers: First Annual Convention, American Society for Quality Control and Second Midwest Quality Control Conference, Hotel Sherman, Chicago, Illinois, Thursday-Friday, June 5 and 6, 1947. Rochester 8, N. Y.: the Society (c/o Alfred L. Davis, Treas., Rochester Institute of Technology), 1947. Pp. vi, 284. Paper, lithotyped. $3.50. For a complete list of the papers and contributors in this volume, *see* 1431-2, 1438–9, 1453–5, 1528, 1562, 1650, 1665, 1691, 1712, 1714, 1718, 1720, 1761, 1767, 1770, 1776, 1781, 1791, 1847–8, 1853, and 1861 in "Statistical Methodology Index, No. 11" in the March 1948 issue of this JOURNAL. *Two reviews follow:*

REVIEW BY H. A. FREEMAN

Associate Professor of Statistics, Massachusetts Institute of Technology
Cambridge, Massachusetts

FIVE papers, three of them by corporation presidents, deal with broad aspects of quality control, such as the views of management and the problems of organization and administration of quality control programs. Seven papers, chiefly by Professors Burr of Purdue and Wescott of Northwestern, describe, quite nonmathematically, the apparatus of elementary statistical quality control. These include quality control charts for continuous and discrete variables, single and double sampling inspection for attributes, and the rudiments of linear correlation. Fifteen papers are devoted to descriptions of successful applications of elementary quality control to particular problems. Some of these deal with subjects to which the methods of statistical quality control are not often applied, for example, personnel data, time study, stockroom inventories, waste and personnel efficiency.

There remain two papers not classified above. One is a very good review of modern methods for variables inspection, including noncentral *t*, by Professor Olds of Carnegie Institute of Technology; from a statistical point of view it is the most interesting paper of all. The second paper, by Professor Peach of North Carolina, discusses the teaching of statistics to engineering students.

These are formative years for organized quality control and this is reflected in the proceedings of the first national conference. As the field begins to establish itself, it relies almost exclusively on a few elementary, easy-to-understand methods, almost all of which stem from the pioneering work of Shewhart, Dodge and others of the Bell Telephone Laboratories. More complicated, though perhaps not more effective, methods will undoubtedly come later. Another feature of this volume is its occasional evangelism. There will probably be less of this at later conferences for it seems certain that quality control methods will ultimately be very generally employed throughout American industry.

Finally, a word to would-be purchasers. This volume should interest anyone who is actually or potentially interested in industrial quality control. Administrators, plant personnel, teachers, and students will all find something in it for them. Its contents will not generally interest professional statisticians, except as they suggest problems which might be attacked by different methods.

The book is produced by the photo offset process and it is paper bound.

REVIEW BY EUGENE L. GRANT

Professor of Economics of Engineering, Stanford University
Stanford, California

ON JUNE 5 and 6, 1947, the second Midwest Quality Control Conference was held in Chicago in combination with the first annual convention of the American Society for Quality Control. Some 38 papers were presented dealing with statistical quality control applications in many different industries including radio and other electrical products, automobile parts, plastics, temperature controls, Scotch tape, pens and pencils, surgical dressings, beer, dairy products, veneer, cleansers, garments, chemicals, ammunition, pharmaceuticals, paper, hybrid corn, and the retail mail-order business. Through the energy of the Proceedings Committee of the Midwest Quality Control Conference, under the chairmanship of Professor Mason E. Wescott of Northwestern University, this volume containing twenty-nine of these papers has been made available to the quality control public. The volume, reflecting as it does current problems and current thinking in the quality control field, should be a useful addition to the library of any industrial statistician or any teacher of industrial statistics or statistical quality control.

A number of the papers are really expositions of principles of statistical method as applied to various industrial situations. These are at different levels of statistical sophistication. A group of five papers by Professor Wescott and Professor Irving W. Burr of Purdue University constitute a presentation of elementary control chart principles and the elements of acceptance sampling. Other papers, calling for greater familiarity with statistical methods on the part of the reader, deal with such topics as the use of modified control limits, acceptance sampling by variables, statistical techniques for research and development, and the fundamentals of correlation.

Most of the papers, however, are descriptions of specific cases or discussions of technical statistical problems or managerial problems arising out of specific cases. The number of industries represented would seem to give an impressive answer to any industrial manager who gives the familiar response, "My business is different," to the suggestion that he look into the applicability of statistical quality control techniques in his own plant. Nevertheless, it should not be concluded from this spread of applications that the time has yet been reached when statistical quality control methods are in widespread use in American industry. In many cases the uses described may be only in a single plant in a large industry. Some of the authors make it clear that they feel that the surface has only been scratched in their organizations. In other papers where this is not specifically stated it is evident that, at least in June 1947 when these papers were presented, the applications were in the beginning stage.

Several of the papers deal with matters somewhat removed from strictly quality control problems. Two papers apply control charts to personnel statistics. Another applies statistical methods to the interpretation of time study data. A most interesting paper of this type is "Scientific Control of Stockroom Inventories" by R. H. Wilson, General Service Manager of Bell Telephone Laboratories; this applies probability mathematics to the determination of ordering points and presents a fresh and stimulating discussion which should be of interest to industrial engineers and cost accountants as well as to statisticians.

The final paper in the volume, "Sampling and Sorting Ammunition for the Attack on Normandy" by Col. Leslie E. Simon, should, in the opinion of this reviewer, become a classic in the quality control field. This is not only because of the spectacular nature of the subject matter but also because of the effective way in which Col. Simon includes good sound advice for quality control men on the nonstatistical matters associated with their jobs.

It is to be hoped that similar volumes of conference papers will be made available from subsequent A.S.Q.C. conventions and subsequent Midwest Quality Control Conferences.

STATISTICAL METHODOLOGY INDEX, NO. 12*

A QUARTERLY GUIDE TO CURRENT LITERATURE

Edited by
OSCAR KRISEN BUROS
Rutgers University

This bibliographic service presents statistical methodology literature (articles, books, theses, and chapters) published in 1945 to date. Anonymous references are listed first; other references follow in alphabetical sequence by first-named authors. Whenever known, the position and address of an author are given the first time his name appears as sole author or as the first listed of two or more authors; otherwise, his position and address are usually repeated for each of his publications. Volume number, issue number (in parentheses), pagination, and date of issue are given for each journal reference. Stars indicate books, theses, and pamphlets; asterisks indicate publications seen by the editor of this INDEX. *Authors of papers on statistical methodology are invited to send reprints to Oscar K. Buros, Rutgers University, New Brunswick, New Jersey in order to facilitate prompt and accurate listings. Foreign-language papers should be accompanied by English translations of titles. Information is desired about references omitted.*

1877. **Batch House Assignable Causes in Density Control.** *Glass Ind* 26(11):512–3+ N '45.*

1878. **New Statistical Method of Predicting Sunspots Aids Radio Propagation Forecasts.** *J Franklin Inst* 244(6):481–7 D '47.*

1879. **Refined Methods of Control of Cordiness and Workability of Glass During Production:** Statistical Control Charts Applied to Daily Measurements of Density. *Glass Ind* 26(2):83 F '45.*

1880. ADAM, W. B. (Deputy Director and Senior Chemist, University of Bristol Research Station, Campden, Glos., England). **The Use of Statistical Methods in Research on Food Canning:** The Application of Statistical Methods to Food Problems. *Analyst* 73(862):7–11 Ja '48.* Abstract *Chem & Ind* (3):43–4 Ja 18 '47.*

1881. ADLER, H. A. (Senior Engineer, Engineering Department, Line Design Division, Adams St., Chicago 90, Ill.), and MILLER, K. W. (Assistant Director, Armour Research Foundation, Illinois Institute of Technology, Chicago 16, Ill.). **A New Approach to Probability Problems in Electrical Engineering.** *Elec Eng Trans Sect* 65(10):630–2 O '46.*

1882. ANDREWS, T. GAYLORD (Scientific Advisor, Research and Development Group, Logistics Division, General Staff, Department of the Army, Washington 25, D. C.; on leave: Assistant Professor of Psychology, University of Chicago, Chicago 37, Ill.). **Statistical Studies in Allergy:** II, A Factorial Analysis. *J Allergy* 19(1):43–6 Ja '48.*

1883. ANTHONY, AUDREE D. (Chemist), and SOUTHER, HOBART (Chief Chemist). (Cone Mills, Greensboro, N. C.) **Quality-Control Charts Aid Carding and Spinning.** *Textile World* 98(3):111–3 Mr '48.*

1884. ASHBY, ERIC (George Harrison Professor of Botany and Director of the Botanical Laboratories and Experimental Grounds, University of Manchester, Manchester, England). **Statistical Ecology:** II, A Reassessment. *Bot R* 14(4):222–34 Ap '48.*

1885. ATHERTON, A. L. (Supervisor of Quality Control, Westinghouse Electric Corp., East Pittsburgh, Pa.) **Lot by Lot Inspection.** *Mat & Meth* 22(6):1792 D '45.*

1886. BAILEY, W. G. (Assistant Actuary of the Eagle Star Insurance Co., Ltd., 38 Moorgate, London E.C.2, England), and HAYCOCKS, H. W. (Secretary, Actuarial Tuition Service, 6 Staple Inn, Holborn,

---

* Numerous references prepared for this issue of the INDEX have been omitted because of space limitations.

London W.C.1, England). **A Synthesis of Methods of Deriving Measures of Decrement From Observed Data.** *J Inst Actuaries* 73(337):179–99, discussion 200–12 D '47.*

1887. BARTON, SAMUEL G. (President, Industrial Surveys Co., Inc., 347 Madison Ave., New York, N. Y.). **Consumer Panels,** pp. 155–63, discussion 170–6. In *Measurement of Consumer Interest* (*see* 1903).

1888. BEALL, GEOFFREY (Research Associate, Institute of Paper Chemistry, Appleton, Wis.). **Quantitative Treatment of Density Distribution Pattern on Photographic Paper as Produced by the Disturbance of a Spotlight Galvanometer.** *Paper Trade J* 124(19):35–7 My 8 '47.* Also in *Paper Ind* 28(12):1801–2 Mr '47.*

1889. BEARD, R. E. (Assistant Fire Manager, Pearl Assurance Co., Ltd., High Holborn, London W.C.1., England). **Some Notes on Approximate Product-Integration.** *J Inst Actuaries* 73(337):356–403, discussion 404–16 D '47.*

1890. BENNETT, B. M. (Research Assistant, Statistical Laboratory), and NEYMAN, J. (Professor of Mathematics and Director of the Statistical Laboratory). (University of California, Berkeley, Calif.) **The Performance Characteristic of Certain Methods for Obtaining Confidence Intervals.** Abstract. *Ann Math Stat* 19(1):116 Mr '48.*

1891. ★BICKERSTAFF, THOMAS ALTON (Professor of Mathematics and Head of the Department, University of Mississippi, University, Miss.). **Certain Order Probabilities in Nonparametric Sampling.** Doctor's thesis, University of Michigan (Ann Arbor, Mich.), 1948.

1892. BICKING, CHARLES A. (Quality Control Engineer, Hercules Powder Co., Wilmington 99, Del.). **Industrial Statistics: Their Place in the Manufacture of Pulp and Paper.** *Pulp & Paper* 49(3):181–8+ '48.*

1893. BICKING, C. A. **Quality Control for the Pulp and Paper Industry.** *Tech Assn Papers* 30:135–9 Je '47.*

1894. BLACKWELL, DAVID (Professor of Mathematics, Howard University, Washington 1, D. C.). **A Renewal Theorem.** *Duke Math J* 15(1):145–50 Mr '48.*

1895. BLANCHE, ERNEST E. (Chief Statistician, Logistics Division, General Staff, U. S. Army, The Pentagon; and Lecturer in Mathematical Statistics, American University; Washington, D. C.). **The Mathematics of Gambling.** *Sch Sci & Math* 46(401):217–27 Mr '46.*

1896. BOCHNER, S. (Professor of Mathematics, Princeton University, Princeton, N. J.). **Finitely Additive Set Functions and Stochastic Processes.** *Proc Nat Acad Sci* 32(10):259–61 O '46.*

1897. BOWIE, O. L. (Mathematician, Watertown Arsenal, Watertown, Mass.). **A Least-Square Application to Relaxation Methods.** *J Appl Physics* 18(9):830–3 S '47.*

1898. BREYER, RALPH F. (Professor of Marketing, Wharton School of Finance and Commerce, University of Pennsylvania, Philadelphia 4, Pa.). **Questions on Specifications for Consumers' Goods From the Point of View of Marketing,** pp. 142–6, discussion 152–4. In *Measurement of Consumer Interest* (*see* 1903).

1899. BRUNK, H. D. (Instructor in Mathematics, The Rice Institute, Houston, Texas). **The Strong Law of Large Numbers.** *Duke Math J* 15(1):181–95 Mr '48.*

1900. BUIST, J. M. (Rubber Service Department), and DAVIES, O. L. (Research Statistician, Dyestuffs Division). (Imperial Chemical Industries, Blackley, Manchester 9, England.) **Statistical Evaluation of Variations in Rubber Processes and Correlation in Physical Properties.** *Rubber Chem & Technol* 20(1):288–300 Ja '47.* Reprinted from *Inst Rubber Ind Trans* 22(2):68–81 Ag '46.*

1901. CATLIN, J. B. (Assistant Technical Director), and STRIEBY, J. G. (Quality Methods Coordinator). (Kimberly-Clark Corporation, Neenah, Wis.) **Method for Evaluating and Reporting Product Control.** *Tech Assn Papers* 30:199–202 Je '47.* Same as 1463.

1902. CHASSAN, JACK (Chief, Health Reports Branch, Office of the Surgeon General, Department of the Army, Washington, D. C.). **Estimating Direct Labor Cost in Aircraft Production.** *Ind Aviation* 3(1):56+ Jl '45.*

1903. CHURCHMAN, C. WEST (Associate Professor of Philosophy), ACKOFF, RUSSELL. (Assistant Professor of Philosophy), (Wayne University, Detroit 1, Mich.); and WAX, MURRAY (Instructor in Philosophy, Temple University, Philadelphia 22, Pa.), Editors. **Measurement of Consumer Interest.** Philadelphia 4, Pa.: University of Pennsylvania Press (3622 Locust St.), 1947. Pp. vi, 214. $3.50. (London E.C.4: Oxford University Press [Amen House, Warwick Square]. 20*s*.)* For contents, *see* 1887, 1898, 1906, 1910–1, 1917, 1943, 1950, 1951.

1904. COHEN, A. C. (Associate Professor of Mathematics, University of Georgia, Athens, Ga.). **Generalized I_n Functions.**

Abstract. *B Am Math Soc* 54(3):287 Mr '48.*

1905. CRAWFORD, ALBERT B. (Director), and BURNHAM, PAUL S. (Associate Director). (Student Appointment Bureau, Yale University, New Haven, Conn.) Chap. 2, **The Measurement of Educational Performance and Basic Statistical Principles,** pp. 21–75. In their *Forecasting College Achievement: A Survey of Aptitude Tests for Higher Education: Part I, General Considerations in the Measurement of Academic Promise.* New Haven 7, Conn.: Yale University Press (143 Elm St.), 1946. Pp. xxi, 291. $3.75.* (London E.C.4: Oxford University Press [Amen House, Warwick Square]. 25*s.*)

1906. CURTISS, J. H. (Chief, National Applied Mathematical Laboratories, National Bureau of Standards, Washington, D. C.). **The Government's Use of Specifications for Consumers' Goods.** pp. 146–9, discussion 152–4. In *Measurement of Consumer Interest* (*see* 1903).

1907. DALTON, A. G. (Superintendent, Quality Control, Western Electric Co., Inc., 195 Broadway, New York, N. Y.). **Some Engineering Aspects of Quality Control.** *Mech Eng* 70(3):205–7+ Mr '48.*

1908. DEEMER, WALTER L., JR. (Lt. Col., U. S. Air Force, 400 Cameron Ave., Chapel Hill, N. C.). **The Power of the t Test and the Estimation of Required Sample Size.** *J Ed Psychol* 38(6):329–42 O '47.*

1909. DEEMER, WALTER L., JR.; HORST, A. P. (Professor of Psychology, University of Washington, Seattle, Wash.); THORNDIKE, R. L. (Professor of Education, Teachers College, Columbia University, New York, N. Y.); and WHITNEY, A. G. (Foreign Service Staff Officer, Division of Foreign Service Personnel, Department of State, Washington, D. C.). Appendix E, **The Correction of Correlation Coefficients for Restriction of Range,** pp. 609–16. In *Records, Analysis, and Test Procedures.* Edited by Walter L. Deemer, Jr. Army Air Forces, Aviation Psychology Program Research Reports, No. 18. Washington, D. C.: Government Printing Office, 1947. Pp. ix, 621. $2.25.*

1910. DEMING, W. EDWARDS (Adviser in Sampling, Bureau of the Budget, Washington, D. C.). **Random Samples and Biased Samples,** pp. 167–70, discussion 170–6. In *Measurement of Consumer Interest* (*see* 1903).

1911. DEMING, W. EDWARDS. **Some Criteria for Judging the Quality of Surveys,** pp. 106–22. In *Measurement of Consumer Interest* (*see* 1903).

1912. DEMING, W. EDWARDS, and SCATES, DOUGLAS E. (Professor of Education, Duke University, Durham, N. C.). **The Need for Statistical Education in High School and College.** *Ed Rec* 29(1):72–80 Ja '48.*

1913. DURAND, DAVID (National Bureau of Economic Research, West 254th St. and Independence Ave., New York 63, N. Y.). **An Appraisal of the Errors Involved in Estimating the Size Distribution of a Given Aggregate Income.** *R Econ & Stat* 30(1):63–8 F '48.*

1914. FINNEY, D. J. (Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford, Oxford, England). **The Adjustment of Biological Assay Results for Variation in Concomitant Observations.** *J Hyg* 45(4):397–406 D '47.*

1915. FINNEY, D. J. **The Principles of Biological Assay.** *Sup J Royal Stat Soc* 9(1–2):46–81, discussion 81–91 '47.*

1916. FINNEY, D. J. **Statistical Science and Agricultural Research.** *Math Gazette* 31(293):21–30 F '47.*

1917. GUTTMAN, LOUIS (Associate Professor of Sociology, Cornell University, Ithaca, N. Y.). **The Cornell Technique for Scale and Intensity Analysis,** pp. 60–84. In *Measurement of Consumer Interest* (*see* 1903).

1918. HASKINS, ELMER E. (Assistant Professor of Mathematics, Air Institute of Technology, Wright Field, Ohio). **Quality Control by Statistical Methods.** *Math Teach* 41(4):172–4 Ap '48.*

1919. HITT, HOMER L. (Professor of Sociology and Rural Sociology and Head of the Department, Louisiana State University, Baton Rouge, La.). **The Use of Selected Cartographic Techniques in Health Research.** *Social Forces* 26(2):189–96 D '47.*

1920. HOCHSTIM, JOSEPH R. (Chief Statistician), and SMITH, DILMAN, M. K. (Vice President). (Opinion Research Corporation, 44 Nassau St., Princeton, N. J.) **Area Sampling or Quota Control?—Three Sampling Experiments.** *Pub Opinion Q* 12(1):73–80 sp '48.*

1921. HOLLEY, J. W. (Graduate Student, Department of Psychology, Northwestern University, Evanston, Ill.). **A Note on the Reflection of Signs in the Extraction of Centroid Factors.** *Psychometrika* 12(4):263–5 D '47.*

1922. HOLZINGER, KARL J. (Professor of Education, University of Chicago, Chicago 37, Ill.). **Factoring Factors.** *J Ed Psychol* 38(6):321–8 O '47.*

1923. HOTELLING, HAROLD (Professor of Mathematical Statistics and Head of the Department; Associate Director, Institute of Statistics; University of North Carolina, Chapel Hill, N. C.). **Points in the Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics.** *Math Teach* 41(2):76–7 F '48.*

1924. JAMES, EDWIN (Associate Professor of Agronomy, University of Georgia, Athens, Ga.). **A Modification of the Formulae for Regression.** *Am Soc Agron J* 39(6): 545–6 Je '47.*

1925. JONES, MELVIN D. (Assistant Professor of Genetics and Plant Breeding, Oklahoma Agricultural and Mechanical College, Stillwater, Okla.). **The Fraction Method of Analysis of Factors: An Aid in Teaching Genetics.** *J Heredity* 38(12): 368–70 D '47.*

1926. JOSEPH, A. W. (Assistant Actuary, Wesleyan and General Assurance Society, Steelhouse Lane, Birmingham 4, England). **Third-Order Moments of a Bivariate Frequency Distribution.** *J Inst Actuaries* 73(337):427–9 D '47.*

1927. JURKAT, ERNEST H. (Director of Research, Wroe Alderson, Alderson & Simons, Inc., 1905 Walnut St., Philadelphia 2, Pa.; and Instructor in Marketing, Graduate School of Business Administration, New York University, 90 Trinity Place, New York, N. Y.). **Statistical Analysis and Interpretation of Chemical Market Research Data.** *Chem Ind* 59(6):1021–4 D '46.*

1928. KIMBALL, BRADFORD F. (Senior Statistician, State Department of Public Service, 233 Broadway, New York 7, N. Y.). **Assignment of Frequencies to a Completely Ordered Set of Sample Data.** *Trans Am Geophysical Union* 27(6):843–6 D '46. Discussion by E. J. GUMBEL (Associate Professor of Mathematics, Brooklyn College, Brooklyn, N. Y.) 28(6):951–2, author's reply 952–3 D '47.*

1929. KOGAN, LEONARD S. (Assistant Director, Institute of Welfare Research, Community Service Society of New York, 105 East 22nd St., New York 10, N. Y.). **Analysis of Variance—Repeated Measurements.** *Psychol B* 45(2):131–43 Mr '48.*

1930. KUZNETS, GEORGE M. (Associate Professor of Agricultural Economics, University of California, Berkeley, Calif.). **The Use of Econometric Models in Agricultural Micro-Economic Studies.** *J Farm Econ* 30(1):131–9, discussion 139–41 F '48.*

1931. KUZNETS, SIMON (Professor of Economic and Social Statistics, University of Pennsylvania, Philadelphia 4, Pa.). **Measurement of Economic Growth,** pp. 10–34. In *The Tasks of Economic History*: Papers Presented at the Seventh Annual Meeting of the Economic History Association, New Haven, Connecticut, September 12–13, 1947. Supplement 7 of the *Journal of Economic History*. New York 3: New York University Press (Washington Square), 1947. Pp. 174. Paper. $2.00.*

1932. LASKER, GABRIEL W. (Assistant Professor of Anatomy, Wayne University, Detroit 26, Mich.). **Penetrance Estimated by the Frequency of Unilateral Occurrences and by Discordance in Monozygotic Twins. With an editorial note by** CHARLES P. WINSOR (Assistant Professor of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.). *Human Biol* 19(4):217–23, note 223–30 D '47.*

1933. LAWSHE, C. H., JR. (Professor of Psychology, Purdue University, Lafayette, Ind.), and MAYER, J. S. (Mayer and Co., Industrial Psychologists, 707 Times-Star Bldg., Cincinnati 2, Ohio). **Studies in Item Analysis: I, The Effect of Two Methods of Item Validation on Test Reliability.** *J Appl Psychol* 31(3):271–7 Je '47.*

1934. LIENAU, C. C. (Associate Professor of Electrical Engineering, Manhattan College, New York, N. Y.). **Quantitative Aspects of Organization.** *Human Biol* 19(4):163–216 D '47.*

1935. MAINLAND, DONALD (Professor of Anatomy, Dalhousie University, Halifax, Nova Scotia). **Statistical Methods in Medical Research: I, Qualitative Statistics (Enumeration Data).** *Can J Res, Sect E Med Sci* 26(1):1–166 F '48.*

1936. ★MASSEY, FRANK J., JR. (Assistant Professor of Mathematics, University of Oregon, Eugene, Ore.). **Estimation of a Distribution Function by Confidence Limits.** Doctor's thesis, University of California (Berkeley, Calif.), 1947.

1937. MAY, KENNETH (Associate Professor of Mathematics, Carleton College, Northfield, Minn.). **Probabilities of Certain Election Results.** *Am Math Mo* 55(4):203–9 Ap '48.*

1938. MAY, KENNETH. **Variation of the Probability of Unfair Election Results: Preliminary Report.** Abstract. *B Am Math Soc* 54(3):288 Mr '48.*

1939. MORLEY, J. F. (*formerly* Head of Development Division); PORRITT, B. D. (*late* Director of Research); and SCOTT, J. R. (Director of Research). (Research Association of British Rubber Manufacturers, 105–7 Lansdowne Road, Croyden, Surrey, England.) **Reproducibility of Tensile, Permanent Set, and Hardness Test**

Results in the Same Laboratory. *Rubber Chem & Technol* 20(2):551–76 Ap '47.*

1940. MOSER, C. A. (Assistant Lecturer in Statistics, London School of Economics and Political Science, University of London, London W.C.2, England). **Statistics and the Citizen.** *Highway* 28:79 F '47.*

1941. NEYMAN, J. (Professor of Mathematics and Director of the Statistical Laboratory), and SCOTT, ELIZABETH L. (Research Assistant and Associate, Department of Mathematics). (University of California, Berkeley 4, Calif.) **Consistent Estimates Based on Partially Consistent Observations.** *Econometrica* 16(1):1–32 Ja '48.*

1942. PAGE, A. (Lecturer in Mathematics, Borough Road College, Isleworth, Middlesex, England). Chap. 12, **Statistics,** pp. 243–74. In his *Algebra.* London E.C.4: University of London Press Ltd. (Warwick Square), 1947. Pp. vii, 346. 18*s*.*

1943. PEACH, PAUL (Associate Professor, Institute of Statistics, North Carolina State College, Raleigh, N. C.). **Laboratory Testing and Actual Quality in the Consumer Goods Field,** pp. 138–42, discussion 152–4. In *Measurement of Consumer Interest* (*see* 1903).

1944. PERKS, WILFRED (Assistant Actuary of the Pearl Assurance Co., Ltd., London W.C.1, England). **Some Observations on Inverse Probability Including a New Indifference Rule.** *J Inst Actuaries* 73(337):285–312; discussion 313–34 D '47.*

1945. PETERSON, DONALD A. (Research Associate, Life Insurance Agency Management Association, 115 Broad St., Hartford 5, Conn.), and GULLIKSEN, HAROLD O. (Professor of Psychology, Princeton University; and Research Director, College Entrance Examination Board; Princeton, N. J.). Appendix E-3, **Construction and Use of Abacs,** pp. 488–505. In *Personnel Research and Test Development in the Bureau of Naval Personnel.* By the Staff, Test and Research Section in Cooperation With N.D.R.C. Project N-106 and the College Entrance Examination Board. Edited by DEWEY B. STUIT. Princeton, N. J.: Princeton University Press, 1947. Pp. xxiv, 513. $7.50.*

1946. RICHARDS, PAUL I. (Scientist, Physics Department, Theoretical Division, Brookhaven National Laboratory, Associated Universities, Inc., Upton, N. Y.). **Probability of Coincidence for Two Periodically Recurring Events.** *Ann Math Stat* 19(1):16–29 Mr '48.*

1947. RYNIN, DAVID (Assistant Professor of Speech, University of California, Berkeley, Calif.). **Probability and Meaning.** *J Philos* 54(22):589–97 O 23 '47.*

1948. SALZER, HERBERT E. (Mathematician, Computation Laboratory, National Bureau of Standards, Washington, D. C.). **Note on Interpolation for a Function of Several Variables.** *B Am Math Soc* 51(4):279–80 Ap '45.*

1949. SALZER, HERBERT E. **Tables of Coefficients for Interpolating in Functions of Two Variables.** *J Math & Physics* 26(4):294–305 Ja '48.*

1950. THURSTONE, L. L. (Charles F. Grey Distinguished Service Professor of Psychology, University of Chicago, Chicago 37, Ill.). **The Prediction of Choice,** pp. 39–55. In *Measurement of Consumer Interest* (*see* 1903).

1951. WILKS, S. S. (Professor of Mathematics and Director of the Section of Mathematical Statistics, Princeton University, Princeton, N. J.). **Research on Consumer Products as a Counterpart of Wartime Research,** pp. 135–8, discussion 152–4. In *Measurement of Consumer Interest* (*see* 1903).

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 43 SEPTEMBER 1948 NUMBER 243

ARTICLES

Significance of Inventories in the Current Economic Situation . . . . LOUIS J. PARADISO 361

Measuring Physical Inventories . . . . . . WEIR M. BROWN 377

On the Determination of Sample Sizes in Designing Experiments . . . . MARILYN HARRIS, D. G. HORVITZ AND A. M. MOOD 391

Application of the Theory of Extreme Values in Fracture Problems . . . . BENJAMIN EPSTEIN 403

Statistics in Insurance

IV. Problems with Sampling Procedures for Reserve Valuations . . . . GEORGE C. CAMPBELL 413

V. Recent Developments in Graduation and Interpolation . . . . THOMAS N. E. GREVILLE 428

VI. Sampling Errors in Mortality and Other Statistics in Life Insurance. . . . . DONALD D. CODY 442

The Edge-Marking Method of Analyzing Data . . L. L. THURSTONE 451

The War Production Board's Statistical Reporting Experience, Parts II and III . . . . . DAVID NOVICK AND GEORGE A. STEINER 463

Statistical Methodology Index, No. 13 . . OSCAR KRISEN BUROS 500

BOOK REVIEWS

CHAPIN, F. STUART, *Experimental Designs in Sociological Research* . . . . O. KEMPTHORNE 489

HILL, A. BRADFORD, *Principles of Medical Statistics, Fourth Edition* . . . . J. YERUSHALMY 492

Hotelling, Harold; Bartky, Walter; Deming, W. Edwards; Friedman, Milton; and Hoel, Paul, *The Teaching of Statistics*, A Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics . . . . Truman L. Kelley 493

Pearson, E. S., chairman; R. G. D. Allen, H. Campion, William Elderton, C. Oswald George, R. F. George, M. Greenwood, D. Heron, J. O. Irwin, H. Leak, E. C. Rhodes, R. Stone, L. H. C. Tippett, and J. Wishart, *The Teaching of Statistics in Universities and Colleges*, based upon a report by the Royal Statistical Society's Committee on the Teaching of Statistics . . . . Truman L. Kelley 493

Levens, A. S., *Nomography* . . . . M. Kac 496

Samuelson, Paul Anthony, *Foundations of Economic Analysis* . . . . Gerhard Tintner 497

Index to Volumes 1–34, 1888–1939, may be obtained from the ASA. The Journal is also indexed in the Industrial Arts Index, the International Index to Periodicals, and the Public Affairs Information Service Bulletin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 243 SEPTEMBER 1948 *Volume 43*

SIGNIFICANCE OF INVENTORIES IN THE CURRENT ECONOMIC SITUATION

LOUIS J. PARADISO*
The Econometric Institute, Inc.

It is the purpose of this paper to: (1) appraise the current inventory level and recent additions in relation to other investment items in the economy and to the total value of output; and (2) analyze the current inventory levels relative to activity by broad industry groups.

SINCE V-J Day, the course of inventory changes has led at times to optimism and at other times to apprehension regarding the future prospects of business. The inventory accumulation which occurred in the last half of 1945, for example, was salutary, since it reflected the gradual filling of the civilian pipelines which were depleted during the war years, and foreshadowed increased production of civilian goods. The rapid accumulation of inventories which occurred in the second half of 1946, on the other hand, was interpreted by many observers as a sign of impending decline in business activity.

Those who are influenced in their business outlook by the course of inventory changes fail to assign to inventories their proper role in the business cycle. In theory, the part which inventories play in business fluctuations has been extensively developed. But theory is often cast aside when inventory accumulation reaches unduly large proportions. It is then argued that since inventory additions are contributing significantly to the total value of output and since such a high rate of accumulation cannot go on indefinitely, that when the inventory accumulation stops, business activity must turn down unless the inventory loss is offset by rising activities elsewhere in the economy. This reasoning overlooks the fact that inventory changes are a symptom and not a prime cause of business fluctuations. The inventory position at the beginning of a change in the direction of the cycle can moderate or accentuate the movement; it can spell profits or loss for an individual firm; and it can prolong or shorten the cyclical swing. In this sense, the appraisal of the trends and levels of inventories by kinds of business and by types of goods is of vital importance to business men as a guide to inventory policy. Inventory trends, however, are not the guides to use in forecasting the turning points of a cycle since the historical records indicate that they tend to lag behind a number of other basic economic factors in the timing of the turning points.

---

* Paper delivered at the 107th Annual Meeting of The American Statistical Association in New York City on December 28, 1948.

THE MAGNITUDE OF RECENT INVENTORY CHANGES IN THE ECONOMY

Although inventory accumulation proceeded at a rapid rate soon after V-J Day, it was not until the second half of 1946, following the end of price controls that the rate of accumulation became so great as to cause considerable concern. During the last quarter of 1946, business added a physical volume of inventories in the prices current at the time amounting to the annual rate of 5.5 billion dollars. In book value, this represented an addition of an annual rate of 16 billion dollars. Thus, the current value of additions to inventories in that quarter accounted for nearly one-fifth of the total gross private domestic investment. Furthermore, this rate of accumulation meant that nearly 5 per cent of all goods (excluding services) produced in the economy during that quarter were for additions to existing stocks.

In ordinary times such a rise would be symptomatic of trouble brewing elsewhere in the economy, such as in reduced buying power of individuals and in distorted price, wage, and profit structures. Analysis of the basic economic developments in the past year, however, shows that demands for all types of goods were on the increase, that while distortions were appearing in the price structure, the problem businessmen were facing was to get a balance in inventories and to get enough goods of the right quality, style, and color to satisfy the ever-growing demands of consumers and producers, both domestic and foreign. The high rate of inventory accumulation experienced in the fourth quarter of 1946 did not continue; nor did this spell recession. As better balance was obtained in the inventory holdings, supplies of goods flowing to ultimate consumers increased and the reduction in the rate of accumulation from 5.5 billion dollars in the fourth quarter of 1946 to less than one billion dollars in the fourth quarter of 1947, as shown in Table I, was offset by rising consumer expenditures, net foreign balance, and producers' durables. The point commonly overlooked is that even though industrial production failed to rise materially during most of 1947, more goods were flowing to the ultimate consumer than in 1946 as less goods were being channeled into inventory additions.

Let us examine in some detail the inventory position for the major groups of industries—manufacturing, wholesale, and retail trades.

METHODS OF ANALYZING INVENTORIES

The usual method businessmen employ in appraising the inventory position is to compare the current stock sales ratio with that in a period which would be considered as normal. This might be the average of a pre-war period of years of good business, or 1940 or 1941. This method has the advantage of simplicity, is easily understood and takes little time to develop. It is useful in many cases where data are not available over a long period and a more refined analysis cannot be made and where the current ratios are substantially different from the period chosen as normal.

TABLE I

	1940	1941	Annual rates from 4th Quarter 1946	4th Quarter 1947
		(Billions of Dollars)		
Change in Inventories:				
Current Valuation	2.3	3.9	5.5	0.6
Book Value	2.4	7.2	16.0	8.0
Gross Private Domestic Investment:	13.0	17.2	30.4	29.9
Gross National Product:	100 5	125.3	218.6	240.9
Ratios:		(*Percentages*)		
Change in Inventories to Gross Private Investment:				
Current Value	17.7	22.7	18.1	2.0
Book Value	18.5	41.9	52.6	26.8
Change in Inventories to Gross National Product:				
Current Value	2.3	3.1	2.5	0.2
Book Value	2.4	5.7	7.3	3.3

Source: U. S. Department of Commerce.

The stock-sales ratio, however, has two distinct disadvantages and on that account can be misleading and result in a wrong and costly policy for buyers and merchandisers. First, the stability of the stock-sales ratio assumes that a given percentage change in sales is accompanied by a similar percentage change in stocks. For example, those who use the stock-sales ratio reason as follows: if it is assumed that a stock-sales ratio of 3 is considered normal in a year such as 1941, then a ratio of 2 in 1947, under conditions of substantially greater sales would imply thinness of inventories. This conclusion may or may not be the case. Further analysis might show that in the particular business a change of ten per cent in sales is accompanied by a change of only 5 per cent in inventories and therefore with greater sales, the stock-sales ratio should be lower in order that stocks be just adequate to support the larger volume of sales. In other words, in this example, the stock-sales ratio of 2 would be in line with the historical experience of the firm and would indicate a balance rather than thinness in stocks. Second, use of the stock-sales ratio assumes that the stocks at any period of time must bear a fixed relation to the rate of sales in that period. Actually, inventory changes are the resultant of dynamic factors and are influenced by changes in activity of previous periods as well as long-term changes in turnover and merchandising policies of business firms.

A more fruitful and logical technique for analyzing inventories is that of developing the relationship of inventories to the basic economic factors affecting their fluctuations. This is the approach used in this paper.

THE DATA

Overall inventory data are not available on a monthly basis before 1939. Year-end data therefore must be relied upon for the earlier years. Even so, the existing data are not complete for all years. Census information is available for some years; data on department store stocks for all years back to 1919; the Bureau of Internal Revenue is the source for value of inventories held by corporations. In this report, the estimates made by Simon Kuznets and the National Bureau of Economic Research have been relied on primarily for the data before 1939. These estimates have been spliced to the Department of Commerce data beginning with December 31, 1938. The benchmark used by the Department was the Census of Manufactures and Census of Business for 1939 which contain data on the value of inventories at the beginning and the end of 1939. Since 1939 monthly data have been compiled by the Department on the basis of a representative sample of firms covering manufacturers, wholesalers and retailers.

In general, the values of inventories analyzed are book value figures which reflect not only the current value of physical additions but also revaluations resulting from price changes. In view of the fact that price changes also affect sales or the value of output, the book value is the more appropriate measure to relate to the value of output or sales.

It is difficult to appraise the reliability of existing data on inventories. Book value figures reflect the various practices of firms which are not uniform from one industry to another or even within an industry. Various tests have indicated that, in general, existing sources tend to provide less reliable data on the *levels* of inventories but are much more accurate with regard to the movements or changes. Since the analysis presented in this report is concerned with the development of historical relationships involving two or more variables, the accuracy of the level of inventories is not as important as the reliability of the size and direction of the inventory change.

MANUFACTURERS' INVENTORIES IN RELATION TO PRODUCTION

It has been found that the value of manufacturers' inventories is dependent not only on the value of output during a given period but also on the value of output in a previous period and on a gradual trend of the residuals of inventories after eliminating the influence of changes in the value of output.

Using the data for the years 1920 to 1941, the regression of inventories on value of products is as follows:

(1) $v_i = 3.45 - .0122t + .0661v_c + .0633v_p,$

where

v_i = average value of manufacturers' inventories (arithmetic average of values at end of current and previous years),
$t = 2(\text{year}-1931) + 1,$
v_c = value of manufactured products in current year,
v_p = value of manufactured products in previous year.

The values are in billions of dollars.

Table II and Chart I compare the actual inventories with the values calculated from the value of product according to regression (1). The fit is obviously very good. For the period 1920–41 the average percentage errors (ratio of the residuals to the calculated values) is only 2.5 per cent; with a maximum error of 8 per cent in 1920; the multiple correlation coefficient is 97 per cent.[1]

The fact that the inventory changes can be estimated very closely from changes in the value of output in the previous and current year provides a standard or a gauge for appraising the current inventory position relative to the value of production.

The formula implies the following three results based on the experience of the pre-war period, 1920–1941:

1. Everything else being equal, the value of manufacturers' inventories declined on the average by 24 million dollars per year. This has resulted from a steady improvement in the flow of materials and from increased efficiency and better planning on the part of management, particularly the gradual introduction of inventory control systems whereby the inventory units are geared to the rate of production.

2. Other factors being equal, an increase (or decrease) of 1 billion dollars in the value of output in the current year has been asso-

---

[1] If the value of output of the previous year is excluded from the analysis, the multiple correlation coefficient is 91 per cent.

ciated with a rise (or decline) of 660 million dollars in the average value of inventories.

3. Assuming the same current year value of output and no

TABLE II

Value of Inventories of Manufacturing Industries

	Total		Non-durables		Calculated Inventories		Per Cent of Residual to Calculated	
	Value of Output*	Average Inventories†	Value of Output‡	Average Inventories†	Total§	Non-durables¶	Total	Non-durables
			Billions of Dollars					
1920	69.5	11.2	33.5	5.8	12.1	6.1	− 8.0	− 4.9
1921	41.6	10.3	26.6	5.2	10.8	5.4	− 4.8	− 3.9
1922	50.0	9.8	28.0	5.1	9.6	5.2	2.0	0
1923	58.2	11.0	32.9	5.7	10.6	5.6	3.7	1.8
1924	53.5	11.5	30.5	5.8	10.8	5.6	6.7	3.6
1925	60.8	11.6	35.1	5.9	11.0	5.8	5.7	1.7
1926	62.5	11.8	36.0	6.2	11.5	6.1	2.5	1.6
1927	60.3	11.5	35.8	6.1	11.5	6.1	0	0
1928	61.2	11.3	36.5	6.2	11.4	6.1	0	1.6
1929	68.0	11.7	38.4	6.4	11.8	6.3	− .9	1.6
1930	51.0	11.2	32.5	6.0	11.1	5.9	.9	1.7
1931	39.8	9.5	26.0	5.1	9.3	5.1	2.1	0
1932	28.5	7.6	17.5	4.2	7.8	4.2	− 2.6	0
1933	30.6	7.2	21.0	4.0	7.2	4.1	0	− 2.4
1934	37.0	7.8	24.0	4.5	7.8	4.4	0	2.3
1935	45.0	8.2	28.1	4.7	8.6	4.8	− 4.9	− 2.1
1936	53.0	9.2	31.0	5.1	9.7	5.2	− 5.4	− 1.9
1937	60.7	10.6	34.8	5.7	10.6	5.3	0	7.5
1938	47.0	10.5	31.5	5.6	10.2	5.5	2.9	1.8
1939	56.8	10.2	33.6	5.5	10.0	5.5	2.0	0
1940	66.0	11.2	36.4	5.8	11.2	5.8	0	0
1941	93.5	13.5	47.4	6.5	13.6	6.7	− 0.7	− 3.0
1942	121.2	17.0	56.4	7.8	17.1	7.9	− 0.6	− 1.3
1943	148.5	17.6	62.3	7.9	20.6	8.7	−17.0	− 9.2
1944	156.3	17.3	66.9	7.8	22.8	9.3	−31.8	−15.1
1945	140.2	16.4	69.1	8.1	22.3	9.7	−36.0	−16.5
1946	125.7	18.3	74.1	9.3	20.2	10.1	10.4	− 6.9
1947	168.8	22.1	95.2	10.8	22.2	11.9	0.5	− 9.2

* Source: 1919–1939, census years from Dept. of Commerce. 1939–1947 from Dept. of Commerce. For inter-census years interpolated, see Appendix—Table 1.

† Source: 1929–1947 from Dept. of Commerce. 1919–1929 from National Bureau of Economic Research.

‡ See Appendix—Table 2.

§ Calculated from Average Value of Manufacturing Inventories $=3.45-.0122t+.0661$ (value of current year in billions of dollars) $+.0633$ (Value of Product previous year in billions of dollars) where $t=2$ (year—1931) $+1$.

¶ Calculated from Average Value of Non-durable Goods Inventories $=1.67-.00986t+.0756$ (current year value of product billions of dollars) $+.0460$ (preceding year value of product billions of dollars), where $t=2$ (year—1931) $+1$.

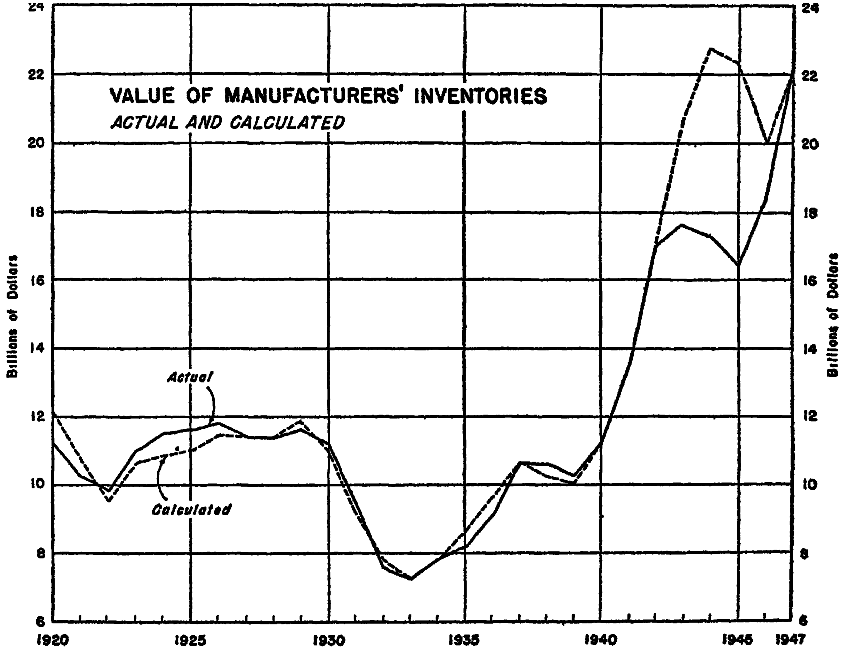
Note: 1939–1947 the value of output is represented by the value of shipments which differs only by small amount from the former.

change in other factors, if the value of output in the previous year had been 1 billion dollars greater (or smaller), the average inventories in the current year would have been 630 million dollars greater (or smaller).

The fact that the value of output of the previous year is about equal in importance to that of the current year in its influence on inventories implies that the current level of inventories is determined approximately by the rate of the value of output six months earlier. This amount of lag in inventory relative to sales is applicable to manufacturing industries as a group and undoubtedly varies industry by industry and company by company.

What is the current position of manufacturers' inventories relative to the pre-war relation of inventories to production? On the basis of the 1947 value of output of 169 billion dollars and that of 1946 of 126 billion dollars, the average value of inventories in 1947 is calculated to be 22.2 billion dollars according to formula (1). The actual average for the year is estimated at 22.1 billion dollars, or about equal to the amount warranted by the value of output. Thus manufacturers' inventories were in line in relation to the pre-war stocks-production relation during

CHART I



1947.[2] The rise in the value of these inventories of about 500 million dollars which occurred during the final quarter of 1947 did not bring stocks out of line with the value of products since the latter also increased from the previous quarter. From September to October, for example, the value of products increased by 8 per cent while the value of inventories rose by less than 1.5 per cent. Thus the value of manufacturers' inventories at the close of 1947 were in balance on the basis of the pre-war stocks-output relations. This condition, however, varies by industries within the manufacturing group.

When manufacturers' inventories are analyzed by the two major classifications of industries—durable and non-durable goods—it is found that in the case of non-durable goods the value of inventories in 1947 was deficient relative to the value of product to the extent of over 1 billion dollars. This is derived from the pre-war relationship of the value of inventories of the non-durable goods industries to value of products of the current year and the preceding year. The formula derived from the data for the period 1920–1941 is as follows for non-durable goods manufacturing industries:

(2) $v_{in} = 1.67 - .00986t + .0756v_c + .046v_p$,

where the symbols have the same meanings as in equation (1), except that v_{in} refers to non-durable goods.

Where $t = 2$ (year—1931)+1—average value of non-durable goods inventories equals arithmetic average of end of current year inventories with end of preceding year.

As Table II and Chart II show, the formula approximates the actual value of inventories very closely, the average error for the period 1920–1941 being 2 per cent, while the index of determination is 98 per cent.[3]

The fact that inventories are currently not up to the levels warranted by production indicates (1) that supplies of certain goods were hard to get and stocks could not be built up to required levels and (2) that manufacturers of non-durables deliberately kept their inventories at a minimum in a period when sales were at record levels.

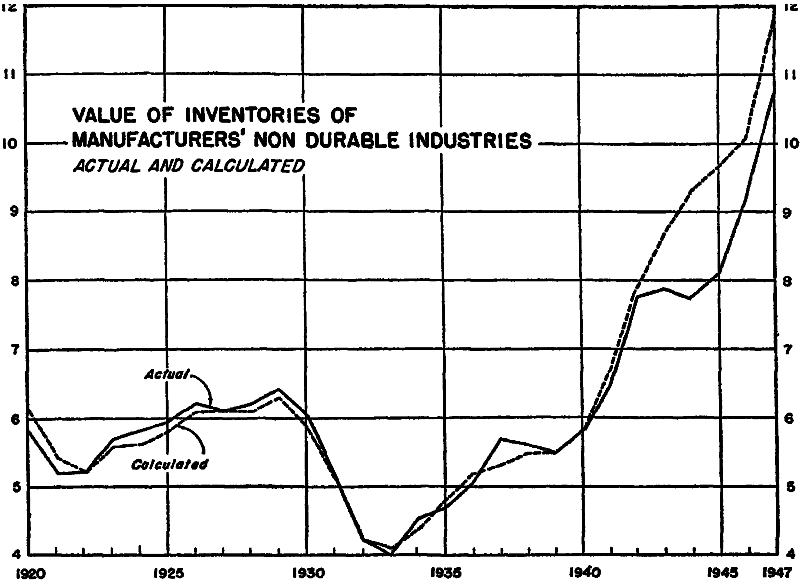
In contrast to the non-durable goods industries, inventories of the durable goods group are currently over one billion dollars above the amount indicated by the pre-war relationship of inventories to value of product of these industries. This is implied from the results obtained

---

[2] It may be noted that in 1940 when inventories were exactly in line with the relationship to value of products as given by formula (1), the stock-sales ratio was .170 whereas in 1947 the ratio was .131. On this basis it would appear that inventories in 1947 were not up to the 1940 position relative to value of products.

[3] If the value of output of the previous year is excluded from the analysis the index of correlation is 95 per cent.

CHART II



by using formulas (1) and (2). At first sight, this would seem to indicate some trouble for these industries. Actually, the reason for the "excessive" inventories in these industries is that there is considerable unbalance in the type of inventory holdings so that many companies hold large stocks of certain parts and components while they lack sufficient supplies of other goods needed to achieve an expanding volume of output. This means that the flow of materials has been uneven and part of the final output was held down because of inability to get a balanced supply of materials going into production. This situation has persisted ever since V-J Day and has been a major factor in limiting the flow of much needed durable goods to the ultimate consumers.

WHOLESALE INVENTORIES

Inventories held by wholesalers are much smaller than those held by retailers or manufacturers in relation to the volume of business done. The ratio of wholesale inventories to wholesale sales in October 1947 was .45; for retail, it was 1.07; and for manufacturing, it was 1.49. Furthermore, the change in inventories is not so great in wholesale for a given change in business as it is in manufacturing or retail. In other words, wholesalers have a more rapid turnover and can get along with relatively smaller stocks.

As in the case of manufacturing, the value of wholesale inventories is related to the value of current year's sales, previous year's sales and a time factor.

The specific formula for the period 1930–1941 is as follows:

(3) $v_{iw} = 1.04 - .032t + .0222s_c + .02726s_p$,

where

v_{iw} = average value of wholesale inventories (average of values at the beginning and end of the current year),

$t = 2\,(\text{year}-1936)+1$,

s_c = current year's sales,

s_p = previous year's sales,

and all money figures are in billions of dollars.

This relationship shows that in fact the level of the previous year's sales has a greater weight than that of the current year. As Table III and Chart III show, the average percentage error for the 12 year period is 1.8 per cent; the index of correlation is .999.[4] The 1947 average value of wholesale inventories estimated at 6.5 billion dollars was 0.9 billion dollars lower than the amount calculated on the basis of the relationship to sales. Thus in contrast to manufacturing, it appears that the value of inventories held by wholesalers is below the economically warranted levels assuming the continuation of pre-war relationships.

RETAIL INVENTORIES

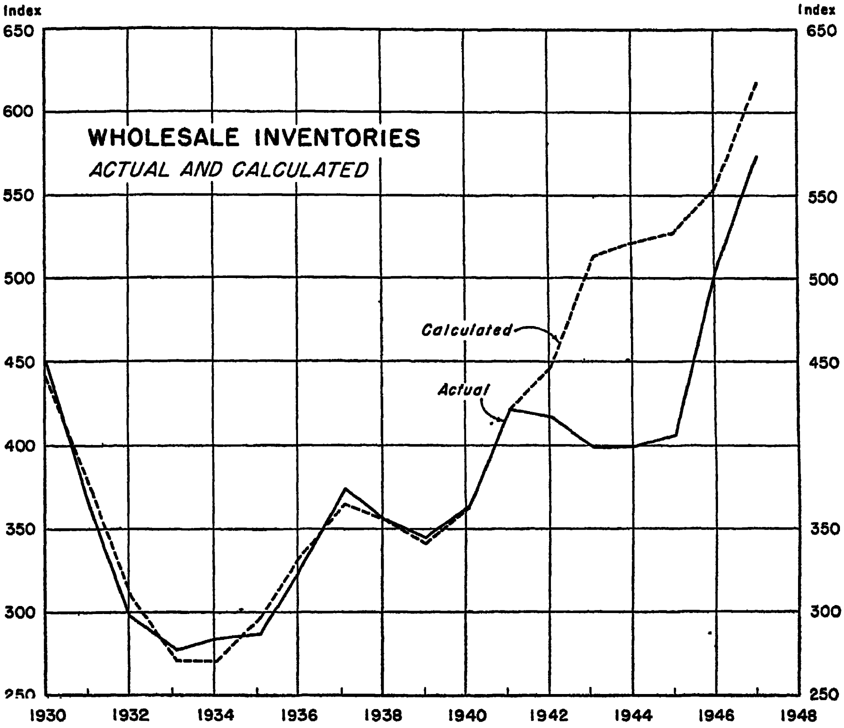
Of especial interest to businessmen is the appraisal of retail inventories since retailers are the first to feel the impact of changing consumer buying, price resistance, and shifting tastes. Many are concerned that retail inventories are rising at uncomfortably rapid rates and that such a development might spell danger in the near future.

At this point it is well to distinguish between the appraisal of retail inventories in the aggregate as against the inventory holdings of particular items by individual firms. In the former case, the impact of excessive or deficient inventories is on the economy as a whole, while in the latter case it affects management's policy decisions to purchase, order, or liquidate goods. In this analysis, we are concerned with the broader aspects of inventories on an industry-wide basis. This will provide retailers with the facts on the over-all inventory position of their industry. It also provides a technique and method which each retailer can apply to analyze the inventory-sales problem for his particular commodities.

---

[4] If the value of output of the previous year is excluded from the analysis the index of correlation is .93.

CHART III



The average value of retail inventories during a year is related to retail sales of the previous year. In addition, after allowing for the level of sales, the value of retail inventories has shown a declining trend in the period 1920–1941. Specifically, the regression equation relating inventories to sales for the period 1920–1941 is as follows:

(4) $v_{ir} = 6.5 - .638t + .815s_c + .507s_p$,

where

v_{ir} = average value of retail inventories (averages of values at the end of the current and previous years),

$t = 2$ (year–1931) $+1$,

s_c = retail sales in the current year,

s_p = retail sales in the previous year.

Values and sales are in billions of dollars.

where $t = 2$(year—1931)$+1$; average value of inventories is the arithmetic average of end of current year and end of previous year inventories.

TABLE III

TRADE INVENTORIES

	Retail Trade		Wholesale*		Calculated Inventories		Per Cent of Residual to Calculated	
	Sales*	Value of Inventories†	Sales*	Value of Inventories†	Retail‡	Wholesale§	Retail	Wholesale
	Billions of Dollars							
1920	43.0	7.45			7.38		.9	
1921	33.8	6.78			6.80		− .3	
1922	34.3	6.04			6.25		− 1.8	
1923	39.3	6.48			6.55		− 1.1	
1924	39.4	6.83			6.69		2.1	
1925	42.9	6.97			6.85		1.8	
1926	45.7	7.19			7.13		.8	
1927	44.8	7.06			7.07		− .1	
1928	46.0	6.96			6.99		− .4	
1929	48.5	7.15			7.13		.3	
1930	42.0	6.61	53.7	4.48	6.60	4.41	.2	1.7
1931	34.8	5.40	40.6	3.64	5.55	3.70	− 2.7	− 1.8
1932	25.0	4.29	30.8	2.95	4.26	3.06	.7	- 3.5
1933	24.5	3.81	30.0	2.74	3.60	2.71	5.8	1.2
1934	28.7	3.91	33.4	2.84	3.79	2.70	3.2	5.2
1935	32.8	4.09	42.8	2.87	4.20	2.94	− 2.6	− 2.5
1936	38.3	4.50	52.2	3.26	4.74	3.34	− 5.1	− 2.5
1937	42.2	4.97	57.6	3.72	5.20	3.65	− 4.4	1.8
1938	38.1	5.06	50.0	3.56	4.94	3.56	− 2.4	0
1939	42.0	5.00	55.3	3.43	4.92	3.41	1.6	0.5
1940	46.4	5.34	61.8	3.64	5.35	3.64	− .2	0
1941	55.5	6.29	83.6	4.21	6.19	4.23	1.6	0.5
1942	57.6	6.94	93.2	4.34	6.69	4.98	3.7	−12.9
1943	63.7	6.69	99.3	3.98	7.17	5.31	6.7	−25.0
1944	69.5	6.44	103.4	3.98	7.82	5.50	−17.6	−27.6
1945	76.6	6.36	105.4	4.14	8.57	5.60	−25.8	−26.1
1946	96.8	7.83	131.6	5.11	10.45	6.17	−25.1	−17.2
1947	109.5	10.09	158.4	6.50	12.38	7.42	−18.5	−12.4

* Source: Department of Commerce 1929–1947. 1919–1929 based on interpolation with consumers, expenditures data on goods, department store sales, etc.

† Source: 1929–1947 from Dept. of Commerce. 1919–1929 from National Bureau of Economic Research.

‡ Calculated from Average Value of Retail Inventories $=6.5-.638t+815$ (retail sales current year bills. of \$) $+.507$ (retail sales previous year bills. of \$), where $t=2$ (year—1931) $+1$.

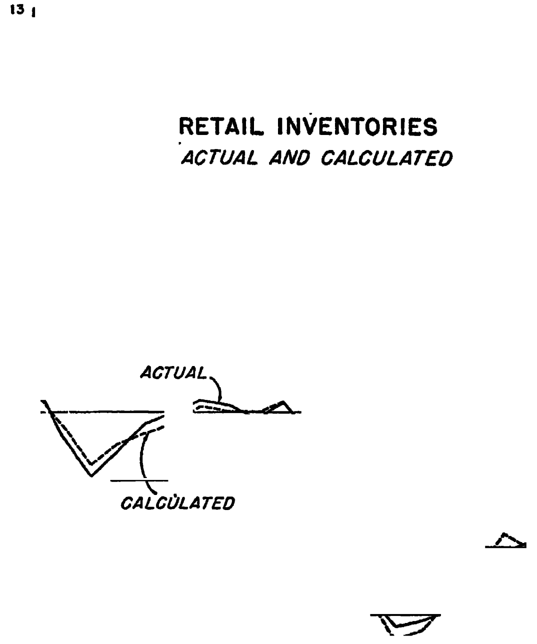
§ Calculated from average value of Wholesale Inventories $=1.04-.032t+.0222$ (current year sales bills. of \$) $+.02726$ (previous year's sales bills. of \$), where $t=2$ (year—1936) $+1$.

This relationship shows that the level of current year retail sales has over one and a half times as much weight as the retail sales of the previous year in its effect on inventories. As Table III and Chart IV show, the correlation is very tight for the period with an average annual error of 2 per cent for the 22-year period, and a maximum residual of only 6 per cent in 1933. The index of correlation is .99.[5]

---

[5] If the value of output is excluded from the analysis the index of correlation is 95 per cent.

CHART IV



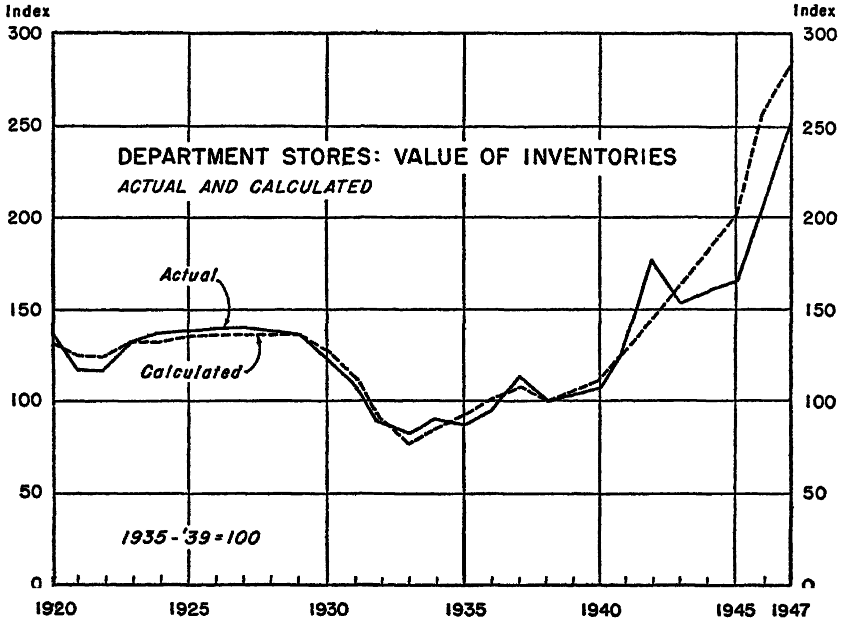
In sharp contrast to the inventory position in manufacturing, retail inventories currently are significantly below the amount calculated from sales on the basis of formula (4). The average value of retail inventories in 1947 was 10 billion dollars whereas on the basis of 1947 retail sales of 109.5 billion dollars and of 1946 sales of 96.8 billion dollars, formula (4) yields a calculated value of inventories of 12.4 billion dollars. This means that retailers today are operating with a significantly smaller inventory than is indicated by the pre-war relation of sales to inventories. There are three reasons for this: first, many goods are still short in supply and retailers cannot accumulate sufficient quantities to bring stocks up to the pre-war relation of stocks to sales. Examples of relative shortages in the consumer durable goods field are of course well-known. To some extent this situation also prevails in certain soft goods. One of the large department stores in New York, for example, found recently that of the customers who failed to buy anything at all, a significant proportion wanted goods which the store did not have on hand. In other words, if certain goods had been more plentiful, sales of many retail stores would have been greater.

Second, many retailers have been pursuing a deliberate policy of holding inventories down to a minimum. They are uncertain of the future because they have never experienced such high levels of sales and they do not want to be caught with high-priced stocks when competition begins in earnest. Third, the economy since V-J Day has been operating under conditions of relatively full employment with a concomitant high level of retail sales. A characteristic feature of this period was the shortages of many goods relative to demand. Competition has only been a factor in recent months. Since retailers could sell practically anything they had, the turnover rate has been greater than in the pre-war period and they did not find it necessary to accumulate inventories to conform with pre-war stock-sales relation in order to achieve a high rate of sales.

DEPARTMENT STORE STOCKS

The data on inventories by kinds of retail business are not sufficiently reliable over a long enough period to permit a correlation analysis with sales. However, analysis of the sales-inventory relation for department stores is probably typical of other major types of non-durable goods stores.

CHART V



APPENDIX—TABLE 1

Method for Interpolating Value of Products in Manufacturing

	(1) Value of Products	(2) FRB Manufacturing Production	(3) BLS Prices Other Than Farmer	(4) Index of Manufacturers' Products
Year	(Bills. Dols.)	(1935–39 = 100)		
1919	60.2	72	161.6	116.4
1920	69.5	74	190.1	140.7
1921	41.6	56	122.9	68.8
1922	50.0	74	119.5	88.4
1923	58.2	86	123.9	106.6
1924	53.5	81	119.2	96.6
1925	60.8	90	124.5	112.0
1926	62.5	95	122.8	116.7
1927	60.3	94	116.2	109.2
1928	61.2	99	116.4	115.2
1929	68.0	110	114.6	126.1
1930	51.0	90	105.5	95.0
1931	39.8	74	91.6	67.8
1932	28.5	57	83.9	47.8
1933	30.6	68	84.7	57.6
1934	37.0	74	94.4	69.9
1935	45.0	87	98.5	85.7
1936	53.0	104	99.1	103.1
1937	60.7	113	105.8	119.6
1938	47.0	87	99.0	86.1
1939	56.8	109	97.6	106.4

Note: Value of Products in odd years from Census; even years interpolated by regression of value of manufacturer's products of Column 4 and census value.

The value of inventories held by department stores is related to the current year sales of department stores and sales of the previous year according to the following regression developed for the period 1920–1941:[6]

(5) $v_{id}=1.4-.993t+.8864s_c+.2539s_p$,
where

v_{id} = average value of inventories of department stores relative to 1935–39 = 100,
$t=2\,(\text{year}-1931)+1$,
s_c = department stores' sales in the current year (1935–39 = 100),
s_p = department stores' sales in the previous year (1935–39 = 100).

Here again, as Chart V shows, the degree of correspondence is very close with an annual average error for the entire period of 2 per cent and an index of correlation of 97.3 per cent. On the basis of an index of department store sales average of 281 for 1947 and 264 for 1946 (1935–39 = 100), the index of the value of department store inventories is calculated to be 285. This compares with the estimated actual index of inventories of 253. Thus, currently, department store stocks are low when appraised in terms of the pre-war stocks-sales relationship. Chief factors in this situation are the deliberate policy of department store executives to hold inventories down to a minimum and inability to get adequate supplies of many goods.

---

[6] The data on inventories and sales are based on published indexes of the Board of Governors of the Federal Reserve System in various bulletins of that agency.

APPENDIX—TABLE 2

Method for Interpolating the Value of Products in Manufacturing Non-durable Goods Industries

Year	(1) Production of Non-durable Goods	(2) Wholesale Prices in Non-durables 1939 = 100	(3) Value of Non-durable Manufacturers	(4) Value of Products
1919	62	186.6	153	36.3
1920	60	218.0	131	33.5
1921	57	139.4	79	26.6
1922	67	132.4	89	28.0
1923	72	135.7	98	32.9
1924	69	132.3	91	30.5
1925	76	140.7	107	35.1
1926	79	137.9	109	36.0
1927	83	128.0	106	35.8
1928	85	129.3	110	36.5
1929	93	126.0	117	38.4
1930	84	114.7	97	32.5
1931	79	97.4	77	26.0
1932	70	86.1	60	18.5
1933	79	87.4	69	21.0
1934	81	98.7	80	24.0
1935	90	105.1	95	28.1
1936	100	106.1	106	31.0
1937	106	112.7	119	34.8
1938	95	101.9	97	31.5
1939	109	100.0	109	33.6

Column 1: Federal Reserve Board.
Column 2: Obtained by combining BLS components of non-durable goods prices by BLS weights.
Column 3: Product of Column 1 and 2.
Column 4: Odd years—Census of Manufactures; even years—interpolated by a regression between Column 3 and census values for the odd years.

MEASURING PHYSICAL INVENTORIES*

WEIR M. BROWN
Board of Governors, Federal Reserve System

SUMMARY

This paper seeks to examine some of the problems encountered in measuring and interpreting the physical volume of inventories and to suggest some methods of approaching these problems. Public attention to the subject of inventory changes certainly has not been lacking in the period since VJ-Day. Indeed, the growth in inventories has to some persons appeared so significant that the period is sometimes characterized simply—probably too simply—as an "inventory boom." This paper will not attempt to describe the course of recent inventory fluctuation, aside from a few isolated illustrative remarks in passing. Rather, this opportunity will be taken to scrutinize the fundamental paths of approach to a field in which existing knowledge must still be regarded, by any sober standard, as limited.

THE TERM "inventories," as ordinarily used, is applied to the stocks of finished goods held by business firms for sale and to stocks of raw or semi-finished materials held either for sale or for processing into marketable goods of another form. Thus in customary usage the term refers principally to stocks in the hands of manufacturers and distributors engaged in making and selling tangible goods in the market. It is not always perceived that inventories in actuality are held also by service and extractive industries, government organizations, and individual consumers, and that conceptually the idea of inventory is universal. In this paper, the term "inventory" is employed in a broad sense to refer to any stock of unconsumed goods, in whatsoever position in the economy it may be held.

If we imagine for a moment a hypothetical economy which is perfectly static—in which industrial techniques and consumer tastes are fixed, the volume of employment and production constant, and with even seasonal variations spirited out of existence—it is obvious that in such a state the amounts and kinds of goods held in inventory likewise would have become perfectly constant, having adjusted themselves to the technical requirements of an unchanging level of consumption of raw materials on the one hand and a fixed production and consumption of finished products on the other.

---

* Paper delivered at the Annual Meeting of the American Statistical Association, New York City, December 28, 1947.

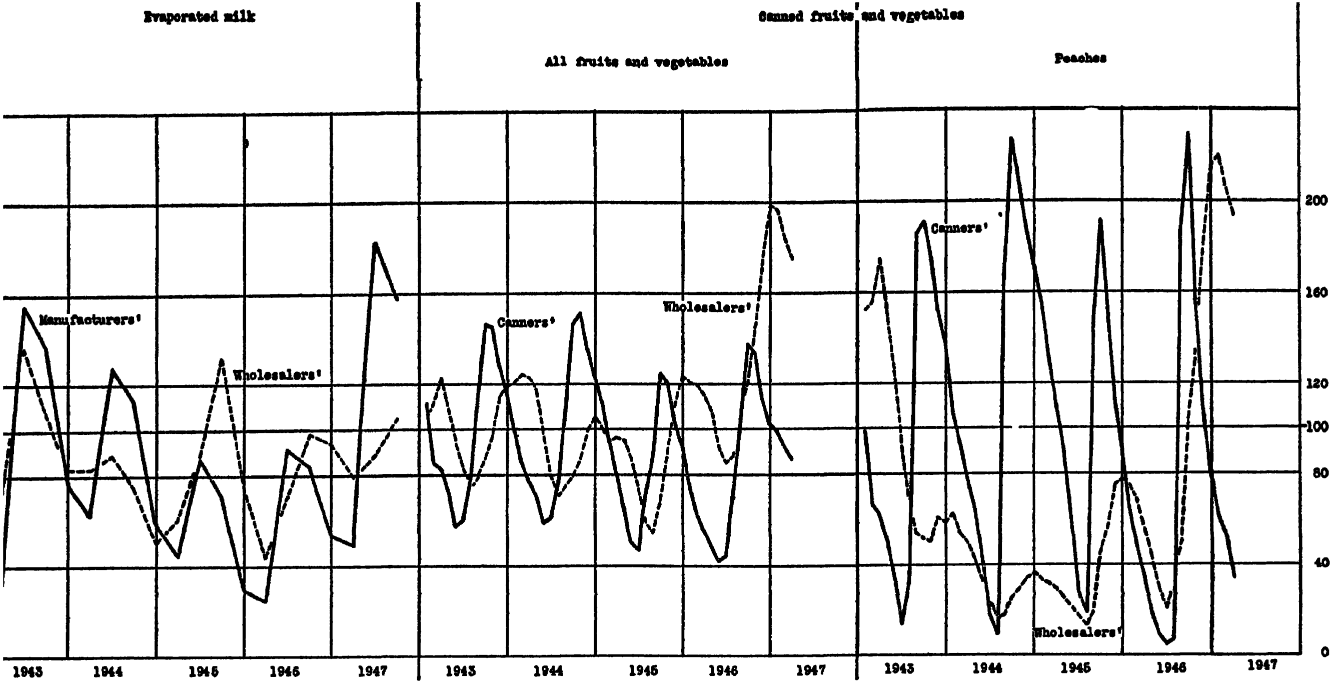
But in the real world, the quantities of goods held in stock in various sectors of the economy are not constant but fluctuate. That is the primary reason why inventories hold any interest for us or present any analytical challenge. An increasing demand for the products of one industry which results in expanded output will ordinarily operate to increase the work-in-process stocks of the industry and may also affect, in one direction or the other, the size of the materials and finished products inventories as well, although the short-run effects (on, say, finished goods stocks) may be quite different from the long-period effects. Of equal importance with changes in demand are varying supply conditions, whether of a periodic or occasional nature. When the market changes represent broad forces affecting the general state of demand or conditions of supply, the shifts imposed on the composition and size of inventory holdings may be very marked and complicated. In turn, the line of influence may flow in the other direction, for inventory levels and movements therein may themselves bear importantly on the economic situation.

When we concentrate on the *change* aspect, it is perhaps easier to elucidate the meaning of inventories and to understand their relevance to the functioning of the system generally. The stock of any good in store at the beginning of a given period constitutes that quantity which was produced in preceding periods but not consumed. It is a capital good (although not necessarily a producer's good, of course). It is analogous in many respects with other kinds of physical assets, both as it relates to the individual holder and in relation to the economy as a whole. A transfer of inventory between two individual firms is an act of investment for the buyer and one of disinvestment for the seller. For the system as a whole, an increase in the total amount of goods held in inventory by all firms constitutes a positive investment, and a corresponding diminution of total stocks of goods represents a consumption of capital or disinvestment. This is a point to which we shall need to return later.

MEASUREMENT CONSIDERED WITH REFERENCE TO PURPOSE

The breadth or narrowness with which inventories are defined depends upon the purpose in view. Moreover, the purpose of the investigation also determines whether stocks will be measured in terms of physical quantities or dollar values, and it influences decisions regarding the exact manner in which any necessary combining of individual series will be accomplished. It was pointed out above that in customary usage the emphasis in discussions of inventories is ordinarily upon so-called "business inventories," that is, the holdings of manufacturing plants and wholesale and retail distributors. There are several respectable reasons in partial justification for this emphasis. For one thing, current statistical data are more readily available for these groups than for some others, although this circumstance is adventitious rather than inherent. A second reason is that the stocks of some other groups which are important quantitatively—such as those of family consumers or of public utility companies—are not being held for sale to others and hence do not "hang over the market," in the same sense as the stocks of manufacturers or retailers. For some purposes, it may be permissible to exclude such holdings by groups other than manufacturers and distributors from one's definition of total stocks; or, if they are not so explicitly excluded, to accept with complacency the present unavailability of the necessary current statistics. Yet for other purposes—and these include the task of analyzing at any time the current business situation—the results will be only partial and probably will be misleading if stocks in all positions are not or cannot be adequately considered. Although it is true that some stocks are not being held for sale and hence do not manifest themselves prominently in direct transactions, such stocks nevertheless do "hang over the market" in a very real sense. The size of these holdings and changes in their magnitudes are important determinants governing the behavior of the holders, influencing not only the strength of their demand for the commodity in question but also conditioning their productive efforts. Manufacturing and trade inventories are undoubtedly large, and attention had been concentrated on them even before Adam Smith used the term "the profits of stock" in describing the activities of the 18th century merchant capitalists. But stocks of unconsumed goods are widely held in many sectors of the economy. An addition to or a withdrawal from stocks represents the amount by which the consumption of a given good differs from the amount produced during the period. There can be no satisfactory measurement of consumption unless adequate information can be developed concerning movements in stocks held in all positions within whatever sector of the economy is under examination.

For some purposes a useful, though seldom used, compilation of inventory data can be made on a commodity basis, tracing the stock holdings of a given commodity or group of commodities through all stages of the productive, distributive, and consumption process. This type of presentation has applications both in the study of particular industries or commodity groups and in describing the inventory posi-



rces: Board of Governors of the Federal Reserve System. Computed from data collected by Department of Agriculture, Bureau of the Census, and National Canners Association. The 1943 average for each series follows: Evaporated milk--mfrs., 241.2 million lbs.; wholesalers, 39.9 million lbs.; all canned fruits & vegetables--canners, 109.1 million cases; wholesalers, 56.3 million cases; canned peaches--canners, 5,279 thousand cases; wholesalers, 2,455 thousand cases. For the fruit & vegetable items, figures for wholesalers include warehouse holdings of retail chain stores.

CHART 1.

COMPARATIVE MOVEMENTS IN PHYSICAL VOLUME OF MANUFACTURERS' AND WHOLESALERS' STOCKS OF SELECTED MANUFACTURED FOODS tion of larger sections of the economy. There are instances in which, for the purpose in hand, it does not matter appreciably at what stage in the economic process stocks of specified finished goods are held, and all such finished stocks can be considered as a total, or at least considered together. This might be the case, for example, if the problem were to ascertain, for certain consumer goods in short supply, what quantity could be available for consumption under conditions of strictly-exercised government controls. Under other circumstances, however, the manner in which a given total amount of inventory were divided between manufacturers, retailers, and other holders might make a substantial difference on the price developments to be anticipated or on the measures required to make a given amount of the commodity available for export. Also, it is scarcely necessary to say that at times when major movements are taking place in prices and the volume of business, the manner in which inventories are divided among holders will help to determine which groups in the economic structure will reap profits or sustain losses.

Chart 1 gives specific illustration to this fact that inventory holdings of the same commodity in different positions may differ considerably, both as to magnitude at any one date and as to movements over time. Stocks in the hands of two groups—manufacturers and wholesale distributors (including chain store warehouses)—are presented for three different manufactured food items: evaporated milk, canned peaches, and total canned fruits and vegetables. The series charted are index numbers representing the physical volume of end-of-month or end-of-quarter stocks during the period 1942–1947, with the 1943 average for each series taken as 100. The several series exhibit different relationships not only in the comparative month-to-month movements of the two groups of holdings but also in their current levels relative to the base year. Since no current data are available on the stocks of these commodities held by retailers, family households, or exporters, the chart does not give a complete account of total stocks in all hands. It does suggest the necessity of caution in interpreting changes in the quantities held in any one position or in drawing inferences from changes in the ratio of stocks to production (or sales) at any one stage.

Another method of classifying or grouping inventories the meaning of which needs to be closely examined with reference to any given purpose is the segregation of stocks according to "stage of fabrication." This is the familiar three-fold classification of raw materials, goods in process, and finished goods. This method of division is a valid one when confined to the various holdings of a particular firm, since such a segregation can be made unambiguous (or virtually so) and since it serves the operational needs of the management. The same method may sometimes be applied to an entire industry, although greater care must be exercised. But if we should desire to classify the total inventories of all manufacturing industry according to stage of fabrication, it is doubtful if our purpose would be accomplished by adding up in one column all the goods classed as "finished" by their producers, in another column all the goods individually reported as "raw materials," etc. Such a process would have the anomalous result of including in the tally of the nation's stocks of finished goods a substantial quantity of wood pulp, sulfuric acid, pig iron, cotton yarn, and other very primary products which, from the standpoint of the economy as a whole, are scarcely finished. It would also place another large portion of the wood pulp, pig iron, and similar stocks in the raw materials category, since they are so regarded by the paper and rayon manufacturers and steel mills which hold them for further processing. An alternative approach to the question of measuring the degree of fabrication of manufacturing stocks which may prove more fruitful, though admittedly more difficult, is that of classifying inventories (measured in either physical or value terms) according to whether they are entirely finished products, intermediate goods, or primary materials with reference to the manufacturing process as a whole.

INVENTORY MEASUREMENT PROBLEMS

We turn now to a more detailed discussion of the measurement problems encountered in dealing with inventories, with special reference to the task of reckoning physical volume. It will be apparent from what was said earlier in analyzing the nature of inventories that they are additive in character, and to a greater degree than some other economic variables, such as the value of product. This characteristic derives from the fact that inventories are pertinent to a point in time rather than to a period of time. A stock of goods is a fund concept rather than a flow concept; it is a quantity rather than a rate. Since the same stock cannot be found in more than one point on the same date, there are no problems of overlapping or duplication (apart from errors in collecting data) encountered in adding up inventories. This is not to suggest, of course, that economic variables which embody flow concepts (i.e., so much quantity per unit of time) cannot be added or that rates cannot be averaged. The problems are greater, however, due to the familiar fact that during any given period a good may pass through more than one manufacturing or distribution stage, and at the end of each stage it will appear as a "product."

Despite this fact that inventories do possess the quality of being additive conceptually, there are nevertheless important problems to be faced in any actual process of summation. These difficulties vary from one situation to another, depending upon the kind of commodity groups involved, upon the extent of the economic area or areas over which combined measurement is sought, and upon the kind of terms (value or quantity) in which the primary data are expressed. Some of the more significant issues can be divided into two broad groups, those arising in the process of dealing with value figures and those arising in the use of quantity figures. Both the physical-volume approach and the dollar-value approach to inventory measurement are large subjects, and in a paper on physical inventory measurement our discussion of value data must be confined largely to the usability of such data in deriving indicators of quantity. First, however, I should like to allude briefly to certain features characteristic of inventory value figures in general which compromise their meaningfulness when aggregated, whether in their original form or in some deflated form.

A fundamental consideration is that figures on inventory values are creatures of corporate cost accounting. As such, they can probably be regarded as having been designed to fit the operational needs of the management—or at least as conforming to whatever practices have grown up and become habitual in the company. But there are several limitations attaching to the collecting and adding together for purposes of economic analysis of inventory figures which represent accounting book records developed for such different purposes by firms utilizing heterogeneous accounting methods. I have reference here principally to the limitations arising out of the great lack of uniformity among companies in the accounting treatment accorded various types of inventory. These characteristics of the original data complicate and equivocate any aggregate measures of the total value of inventories.

But a much more difficult task of measurement and interpretation is encountered if one attempts to employ any given series comprising the total dollar value of inventories as source material from which to derive an indication of movements in the physical quantity of stocks. It is apparent that inventory value figures, like any other value aggregates, are compounds of two types of element—a quantity element and a valuation or pecuniary element—and that changes in the value of inventories may be brought about by changes in either component. Because of the fact that the value of inventories does embody a pecuniary element and is known to respond, to some degree and with more or less flexibility, to changes in the prices at which goods are exchanged in the market, it is often thought that inventories reported in value terms should be deflated by some generally-known price index in an endeavor to obtain a measure of quantity changes apart from the influence of price movements. But the pecuniary elements which are embedded in accounting records of inventory value are not comparable with available price index series. In fact, these pecuniary elements do *not* constitute price data at all. Consequently, an attempted physical-volume determination through a deflation of an inventory value index confronts the obstacles (a) that corporate inventory records usually do not carry stocks of finished goods at current market selling prices but at one or another of several different cost figures having various time references; (b) that a significant portion of total inventories in some sectors of the economy represents intermediate products or goods in process which seldom if ever enter the market; (c) that many firms follow a practice of making certain adjustments in one or more of their inventory figures at the end of an accounting period to reflect changes in replacement costs; (d) that it is difficult to achieve comparable commodity coverage, even for finished goods, between the price index and the inventory value index; and (e) that, even when the commodities covered conform closely, the weighting systems employed in available price indexes take account of relative importance of various goods sold rather than relative importance of different goods in inventory. Because of these difficulties one could hardly expect a process of deflating inventory value data to yield a satisfactory measure of movements in the physical magnitude of stocks.

In a recent investigation of physical stocks in the hands of food manufacturers,[1] the deflation procedure was also tried to see how a series showing the value of inventory in terms of constant prices would compare, in this field, with an index of the physical quantity of stocks measured directly. The differences in behavior of the two types of index were marked. The deflated value series showed much less regularity of seasonal movement within itself than did the physical index; the short-term movements of the two series did not correspond closely; and there were significant differences in the postwar levels of the two indexes relative to the base period. Although the dissimilarity was partly attributable to differences in coverage and differences between the weighting system employed in the price deflator and that of the inventory series, a substantial part of the dissimilarity was believed explainable by the difficulties inherent in attempting to relate the cost-accounting valuations represented in inventory book-value data to changes in market prices.

---

[1] Described in an unpublished manuscript, which has been circulated within the Federal Reserve System in mimeographed form, "Changes in Physical Stocks Held by Food Manufacturers" (December 1947).

DIRECT MEASUREMENT OF PHYSICAL INVENTORIES

In view of the dubious meaning of figures arrived at through a deflation of inventory value series, it seems desirable to approach the task of ascertaining the physical volume of inventories directly through the use of data expressed in quantity terms. This is no novel idea, of course. Physical inventory statistics undoubtedly would be more widely employed if they were more readily available. Some of the problems encountered in developing such measures should be described.

First, as to individual series. Current series on inventories in quantity terms are regularly collected for numerous commodities. There is cause for gratitude in the fact that the number of individual commodities covered by current series has slowly increased. Most of the series emanate from government agencies or trade associations, and ordinarily they relate to items which are the major products of or principal materials consumed in a given industry or trade group. The available stocks series ordinarily cover a very high proportion or the entire quantity of such stocks held in particular positions.

After these encouraging facts have been enumerated, it still must be acknowledged that the series presently available on commodity stocks are inadequate in several respects. (1) Although there are many stocks series on manufacturing industries, a few for agriculture and wholesaling, and a scattering for other groups such as importers and Federal government agencies, the coverage of industries is by no means uniform. Even within the manufacturing sector of the economy, where coverage is greatest, there are wide differences in the amount of physical stocks data available. If we consider the manufacturing field according to Census industry groups, we find reasonably satisfactory current data for foods, textile-mill products, paper and pulp, rubber, petroleum and coal products, and the stone, clay, and glass group. On the other hand, there are no continuous stock series whatsoever for the motor-vehicle, transportation equipment, or furniture groups and virtually none for machinery. The remaining industries fall somewhere between these extremes. (2) A second point of inadequacy is that coverage is not uniform as to type of commodity. More series are available on raw materials or on finished products of a primary sort than on highly advanced or diversified finished products. There are very few series on stock holdings of wrapping and packaging materials or of fuel held by consuming plants. Fortunately, the proportion of total inventories left uncovered by these circumstances is smaller than the absolute number. (3) Outside of the manufacturing sector, series on physical stocks are even less numerous. In general, data on physical stocks are fewer than inventory value data in almost every major sector of the economy. The comparison is especially unfavorable in retail trade. Exceptions to this generalization are found in agriculture and in a few other cases. It is noteworthy that there have even been some attempts to gather information on consumer stocks, though there are few continuous series in this important field.

Apart from the question of obtaining individual series for particular commodities, there are the further problems of how individual series may be combined and how inventory movements in broad areas of the economy may be represented. Some tentative recommendations on these matters can be advanced, based on the investigation we have made of stocks in the hands of food manufacturers, although methods employed in that field might require more or less alteration if applied elsewhere. First of all, where the series to be combined refer to many heterogeneous commodities, group measures cannot be obtained by direct summation, and any index number will have to be an average of relatives. In most cases, some system of explicit weights will appear desirable, in order to accord greater influence in the total index to movements in stocks of major importance than to movements in those of minor importance. In constructing the index of stocks held by food manufacturers, it was decided that the best criterion of "importance" for this purpose was provided by the value of inventory of the respective commodities in 1939 as recorded by the Census of Manufactures. The use of weights derived from inventory value data means that insofar as the "q" elements embodied in the value figures correspond with the "q's" in the series themselves, we are in effect weighting with 1939 prices. In making use of Census of Manufactures data for deriving weights, two other technical problems arise which can only be mentioned in passing. One is the task of deriving from Census inventory value figures collected on an industry basis estimates of inventory values for particular commodities within the industry, and the other is the adjustment required by the fact that the Census figures represent year-end values rather than averages for the year.

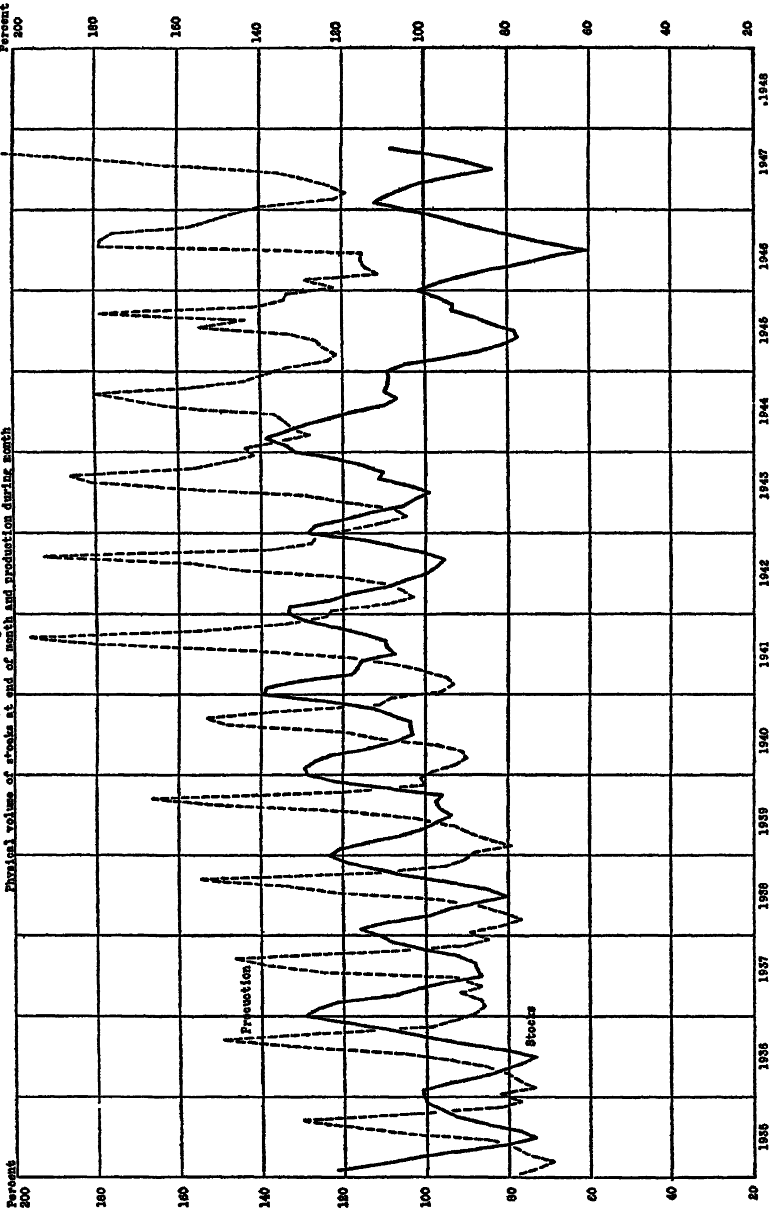
An important fact to be observed is that in measuring stocks by this method many items are left "unrepresented," i.e., any representation is only by the index as a whole and not by parts. The decision to use this method is dictated by the difficulty of inferring movements in the inventory of one commodity from movements in stocks of some other commodity for which data are available. In measuring production, there are certain well-known relationships which can be relied upon with more or less security in statistical measurement. For example, certain goods are made as "joint products" of the same manufacturing process. If a series of figures is available on the slaughter of cattle, for instance, it can be made to serve as a measure not only of packing-house production of beef but also of cattle hide production. Sometimes the consumption of materials is used to represent the activity involved in processing these materials into finished products. There is also a relationship, though of a different sort, between the output of a raw material and that of a finished product in which it is a major ingredient. But relationships between the stocks of two different goods are much less predictable than output relationships. Even in the instance of jointly-produced goods like beef and hides, there can be no assurance that if the packer's stocks of hides are growing his stocks of beef are likewise increasing. Demand factors may be such as to move stocks of the joint product in the opposite direction or to leave them unchanged. There are analogous difficulties in dealing with raw material stocks.

If one of the available series included in an index of stocks were to be accepted as measuring not only the fluctuations of that one commodity but also variations in some other material or product, the importance of the unrepresented items could be assigned, by means of "imputed weights," to the included series. But the decision to impute weights in this fashion would involve the assumption that fluctuations in an available series would faithfully reflect movements in the unrepresented stocks whose weights it carried. In food manufacturing, the available statistical material has not yielded empirically any reliable relationships on the strength of which a system of indirect representation could be justified; moreover, there are reasons *a priori*, as mentioned in the paragraph above, for believing that in some cases, at least, no reliable relationship exists.

For these reasons, it appeared best practice in computing the index of food manufacturers' stocks to let any given series represent only the stocks of the commodities covered directly; the weights of other items were not imputed to the series of represented items. When this method is followed, the selection of items for inclusion in an index is based on their measurability and the resulting index is representative of movements not of all the relevant stocks but only of those for which quantity data are included in the index. In the case of the food index, further investigation is required to establish the validity of using this index as a measure of total food stocks in manufacturing. However, the inference should not be drawn that these uncovered items would actually have been "represented"—in the sense that the index would be more faithfully reflective of movements of all food inventories—if weights

STOCKS AND PRODUCTION OF FOOD MANUFACTURERS
(Without seasonal adjustment)
1935-39 average = 100

Physical volume of stocks at end of month and production during month



Source: Board of Governors of the Federal Reserve System. For comparison with the stock index, some food items were deleted from the published "manufactured foods" production figures. Alcoholic beverages are not included in either series.

attributed to them had been assigned to some one or more series of represented stocks. The imputation of weights, the process of nominating one statistical series to double or act as proxy for another, has many legitimate uses, but it is invalid unless the relationship between the two is so well established that a substitute can be relied upon to behave in a manner similar to, or predictably different from, that of the absent member. Although this has not been found to be the case in the work here described, there is no implication that the same lack of consistent relationship would necessarily be found in stock data in other commodity fields.

CONCLUSIONS

Any conclusions to which this examination may lead should be regarded as tentative. However, I should like, in review, to emphasize certain points. Just as there can be no single, all-purpose index of prices, there is no single measure of inventories—nor even a single method of measurement—which can serve all purposes. We need to be critical in scrutinizing any index or other compilation of inventory data to make sure of its meaning or relevance to the problem in hand. Both the dollar-value approach and the physical-volume approach have an important place in economic analysis; but they do not always have the *same* place. Speaking generally, inventory value figures are more appropriate for use in conjunction with corporate financial-statement data and with some other types of data expressed in dollar values. Physical stocks figures are more appropriate for use in connection with production measures and certain other data reckoned in quantity terms. Because of fundamental differences in their nature, data on book values often cannot be made to yield fully satisfactory indications of changes in the physical quantities of stocks. Physical stocks held in some parts of the economy can be directly measured with a fair degree of reliability from present data, but before indexes representative of very broad areas of the economy can be constructed there must be a substantial increase in the number and coverage of basic series. Since the concept of inventory is considerably broader, in my opinion, than current usage would indicate, there is a corresponding need to extend the investigation of stocks to commodities and to economic groups—notably family households—hitherto largely ignored.

The further study of inventories should also be extended in two other directions. Detailed study is needed of the relationships between the size of inventories in a particular industry and its rate of activity. Investigation to date seems to point to several preliminary conclusions. (1) The ratio of stocks to sales (or production) is not constant at all levels of activity. (2) Different classes of inventory within the same industry exhibit different movements in relation to activity. Often there is a fairly close association between the amount of goods in process and the rate of operations; but there are many possible relationships between changes in activity and movements in stocks of materials or finished goods. (3) The relationship between inventory holdings and sales is influenced by the general state of activity as well as by the level of activity in the industry in question. Stocks are less likely to rise *pari passu* with increases in activity when the economy is operating close to full employment than at lower levels. (4) The degree of correspondence depends not only upon the level from which a given movement in activity or change in stocks commences but also upon its rapidity. (Some of these relationships can be seen by a careful reading of Chart 2, which shows comparative monthly movements of production and stocks of an important group of foods during the period since 1935. The correspondence between the growth of output and of inventories of these goods which had characterized most of the period through 1940 disappeared after that time, and the high levels of food production achieved during the war and subsequent to it were accomplished without a commensurate rise in stocks held at the manufacturing level.)

Another matter requiring further study is the effect on the economy as a whole of various amounts of inventory change. In the imaginary static society earlier referred to, the levels of inventories were constant; there was never a so-called "net change in inventories"; or, rather the net change was always zero. In the actual world, although large net changes in inventories are ordinarily associated with periods of significant shifts in the general economic situation, it would be dangerous to treat a zero net change in inventories as necessarily signifying the existence of stable conditions. The relationships between the general level of activity and the amount of inventory change need to be investigated in terms of gross changes as well as net.

ON THE DETERMINATION OF SAMPLE SIZES IN DESIGNING EXPERIMENTS*

MARILYN HARRIS,† D. G. HORVITZ, AND A. M. MOOD
Iowa State College

Methods are developed for determining sample sizes required to estimate means of normal populations with given precision, or to give significance when means differ by a specified amount. The methods assume that an estimate of the variance has been obtained from a previous or preliminary experiment. A simple device is suggested for using a priori information about the variance when a formal estimate is not available.

1. INTRODUCTION

STATISTICIANS frequently encounter the problem of determining the sample size required to obtain a given result, for example, to estimate a mean by a confidence interval no longer than a specified amount. or to show a significant difference between two means if the difference is greater than a specified amount. The problem cannot be solved if nothing is known of the population distribution, but this is a rather unusual situation. Experimental workers generally know, for example, whether their populations are sensibly normal or not, and if not, what transformations will make them so. But this knowledge (of the form of the population distribution) is still not enough to make the problem determinate. It is necessary also to know something about the order of magnitude of the standard deviation.

The experimental worker ordinarily has information of greater or less precision about the magnitude of the variance. He may have a rather precise estimate based on past experiments with the same population, or he may have only a guess. Usually his information is between these extremes, being based on past experiments with similar but not identical populations. We are concerned with finding a method for using information of this kind in planning an experimental investigation. However, we shall first consider the more special situation in which there exists an estimate of the variance with known precision. Then in the final section of the paper we shall return to the more common situation in which the estimate of the variance is of dubious precision. The technique described there is admittedly crude, but it does provide a method for dealing with an important practical problem.

---

* Most of this work was done under contract N7onr371 with the Office of Naval Research.

† Now at Hunter College.

In the following sections we shall deal with the problem of determining sample sizes when:

1. The population is normal, and
2. An estimate s_1^2 of the variance is available.

Thus we assume at the outset that there is definite information about the variance in the form of an estimate obtained from a previous experiment with the same population. The methods to be presented are appropriate when an experiment is to be repeated once but there is no thought of successive repetition as is the case, for example, in quality control work. We are concerned here with the circumstance in which a research worker has turned up a result which, though not statistically significant, would be important if it were reproduced in a more definitive experiment. Thus a person trying various combinations of materials in an attempt to increase the life of a certain manufactured product may find one which gives an estimated mean life much greater than that of the standard product. The confidence interval for the new mean however, is too broad to give positive evidence of the superiority of the new combination of materials, and it is desired to estimate the mean with greater precision in a second experiment, or to test whether the mean for the new combination exceeds that for the standard by a specified amount in a second experiment.

2. SAMPLE SIZE FOR A SPECIFIED CONFIDENCE INTERVAL

Let us suppose that an experimenter wishes to estimate the mean of a normal population with a .99 confidence interval no longer than four units, and that he has an estimate, $s_1^2=9$, of the population variance with $m=8$ degrees of freedom. Let us further suppose that he can be induced to state his problem in the following manner:

> I have an estimate, nine, of the variance of a normal population with eight degrees of freedom. I wish to draw a sample sufficiently large that the probability will be .95 that the .99 confidence interval will have length less than four units.

It is impossible to guarantee that the confidence interval will be no longer than a specified amount, but one can guarantee it with probability .95 (or any other given probability).

This problem is readily solved with existing tables. Let $n+1$ be the required sample size so that the second estimate s_2^2 of the population variance will have n degrees of freedom. The general relation which determines n (see appendix) is:

$$(d/s_1)^2 = t^2_{1-\alpha}(n)F_{1-\beta}(n, m)/(n + 1) \qquad (1)$$

where the symbols have these meanings: α is the confidence coefficient of the confidence interval, β is the probability that the specified length will not be exceeded, d is half the specified maximum length of the interval, $t_{1-\alpha}(n)$ is the $1-\alpha$ critical level of Student's t for n degrees of freedom, and $F_{1-\beta}(n, m)$ is the $1-\beta$ critical level of the variance ratio for n and m degrees of freedom.

The relation cannot be solved explicitly for n. Probably as good a method as any for finding n is to determine it by trial and error. One might, for example, compute a value V for the right side of (1) using a trial value of n. If V turns out to be larger than $(d/s_1)^2$, a larger value of n would be used in the next trial; otherwise a smaller value of n would be tried. In the specific example above where $s_1=3$, $m=8$, $\alpha=.99$, $\beta=.95$, $d=2$, we may try $n=40$, for example, to get

$$V = (7.31)(3.05)/41 = .54,$$

which is larger than $(d/s_1)^2=.444$, hence a larger value of n is needed. After a few trials n is found to be 48 so that the required sample size is 49.

Critical values of the variance ratio for $\beta=.95$ and .99 are tabulated in almost every statistics book; for other values of β one may consult references 1 and 2 at the end of the paper. The value of t^2 in (1) is conveniently tabulated in the F tables since $F_{1-\alpha}(1, n)=t^2_{1-\alpha}(n)$.

We may observe that the result, $n+1=49$, is considerably larger than what one would have obtained by employing this commonly used argument: The population standard deviation is about 3, hence the mean of a sample of size $n+1$ will have a standard error of $3/\sqrt{n+1}$. For a 99% confidence interval $t_{.01}$ will be 2.8 or 2.9, hence the half confidence interval will have a length of about $(2.9)(3)/\sqrt{n+1}$. On putting this equal to two, one finds $n+1=19$. This argument essentially puts $F=1$ in equation (1) and corresponds roughly to putting $\beta=0.5$; that is, with $n+1=19$ one has about a 50–50 chance that the resulting confidence interval will have a length smaller than four.

3 TESTS OF SIGNIFICANCE OF THE MEAN

The problem of testing the significance of a mean is closely related to the problem of confidence interval estimation of the mean. Let us suppose the particular problem considered in section two is changed to read as follows:

> I have an estimate, $s_1^2=9$, of the population variance with $m=8$ degrees of freedom. I wish to draw a sample large enough that the probability will be $\beta=.80$ that I will get significance at the $\gamma=.05$ level in testing whether the mean exceeds 10 if it actually exceeds 10 by as much as $a=1.5$.

TABLE I

VALUES OF $k=a/s_1$ FOR $\beta=80$
$\gamma=05$ FOR ONE-TAILED TESTS, .10 FOR TWO-TAILED TESTS

m n	1	2	3	4	5	6	8	12	16	24	32	∞
1	13.8	8.52	7.39	6.93	6.68	6.51	6.31	6.13	6.04	5.96	5.92	5.79
2	5.88	3.51	3.02	2.81	2.70	2.62	2.53	2.45	2.41	2.37	2.35	2 30
3	4.30	2.55	2.20	2.03	1.96	1.91	1.85	1.78	1.75	1.72	1.70	1.65
4	3.55	2.10	1.80	1.67	1.60	1.56	1.50	1.45	1.43	1.40	1.39	1.36
5	3.12	1.85	1.58	1.47	1.41	1.37	1.32	1.28	1.25	1.23	1.22	1.18
6	2.81	1.66	1.43	1.32	1.27	1.23	1.19	1.15	1.13	1.11	1.10	1.07
7	2.56	1.52	1.30	1.21	1.16	1.12	1.08	1.05	1.03	1.02	1.01	.979
8	2.37	1.41	1.21	1.12	1.07	1.04	1.00	.972	.956	.940	.932	.910
9	2.23	1.32	1.14	1.05	1.01	.978	.944	.913	.898	.883	.875	.854
10	2.11	1.25	1.07	.993	.952	.925	.893	.863	.849	.835	.828	.805
12	1.92	1.14	.975	.902	.865	.840	.811	.784	.771	.758	.752	.732
14	1.77	1 05	.899	.831	.797	.775	.748	.723	.710	.699	.693	.676
16	1.65	.976	.838	.775	.743	.722	.697	.673	.662	.651	.646	.631
18	1.56	.921	.790	.731	.701	.681	.658	.635	.624	.614	.609	.594
20	1.48	.873	.750	.693	.665	.646	.624	.602	.592	.583	.578	.563
25	1.32	.779	.669	.619	.593	.577	.557	.538	.529	.520	.515	.502
30	1.20	.708	.608	.563	.540	.525	.507	.489	.481	.473	.469	.456
40	1.04	.613	.526	.486	.467	.454	.438	.423	.416	.409	.405	.395
50	.925	.548	.471	.435	.417	.405	.391	.378	371	.365	.362	.353
60	.844	.499	.429	.396	.380	.369	.356	.344	.338	.333	.330	.322
80	.730	.432	.371	.342	.328	.319	.308	.298	.292	.288	.285	.278
100	.652	.385	.331	.306	.293	.285	.275	.266	.261	.257	.255	.249

The size of the specified mean is not relevant to the problem; the critical figure is $a=1.5$. In order that the sample size be determined it is necessary that the experimenter come to a decision about the magnitudes of β, γ, and a.

Some new tables are required to solve this problem, one for each combination of values of γ and β. Two are presented here—one for $\gamma=.05$, $\beta=.80$, and one for $\gamma=.05$, $\beta=.95$. The theoretical basis of the tables is given in the appendix. The number of degrees of freedom n in the second estimate s_2^2 of the variance is found by computing the number

$$k = a/s_1 \tag{2}$$

then locating that value in the column headed by m; the number opposite k in the column on the left is the required value of n. In the example given above, $k=1.5/3=0.5$. This value is found in the column headed by 8 in Table I to correspond to $n=31$, so that the sample size would be 32.

Interpolation in these tables is nearly linear on the reciprocal of m and is also nearly linear on the reciprocal of the square root of n.

TABLE II

VALUES OF $k = a/s_1$ FOR $\beta = .95$
$\gamma = .05$ FOR ONE-TAILED TESTS, .10 FOR TWO-TAILED TESTS

m n	1	2	3	4	5	6	8	12	16	24	32	∞
1	57.1	19.5	14.4	12.6	11.6	11.0	10.4	9.85	9.58	9.33	9.21	8.86
2	24.2	7.74	5.60	4.77	4.39	4.15	3.86	3.61	3.49	3.38	3.33	3.19
3	17.6	5.58	4.03	3.39	3.13	2.94	2.74	2.55	2.46	2.39	2.35	2.23
4	14.5	4.58	3.28	2.79	2.56	2.40	2.23	2.08	2.01	1.94	1.91	1.82
5	12.6	3.97	2.88	2.41	2.23	2.09	1.93	1.82	1.76	1.69	1.66	1.58
6	11.2	3.55	2.57	2.17	2.00	1.88	1.73	1.62	1.57	1.52	1.49	1.42
7	10.3	3.26	2.36	1.99	1.83	1.72	1.58	1.48	1.43	1.38	1.36	1.30
8	9.70	3.05	2.19	1.86	1.70	1.60	1.48	1.39	1.34	1.29	1.27	1.21
9	9.12	2.87	2.06	1.75	1.60	1.50	1.39	1.30	1.26	1.21	1.19	1.13
10	8.62	2.72	1.95	1.65	1.51	1.42	1.32	1.23	1.19	1.15	1.13	1.07
12	7.83	2.47	1.77	1.50	1.37	1.29	1.20	1.12	1.08	1.04	1.02	.971
14	7.22	2.28	1.63	1.38	1.26	1.19	1.11	1.03	.993	.959	.942	.893
16	6.73	2.13	1.52	1.29	1.18	1.11	1.03	.959	.924	.893	.878	.834
18	6.35	2.01	1.44	1.22	1.11	1.04	.972	.904	.872	.842	.828	.785
20	6.02	1.90	1.36	1.15	1.05	.991	.921	.858	.827	.798	.785	.744
25	5.37	1.70	1.22	1.03	.940	.884	.822	.765	.738	.712	.700	.663
30	4.89	1.54	1.11	.935	.855	.804	.748	.695	.671	.647	.636	.605
40	4.23	1.33	.962	.809	.739	.696	.646	.601	.580	.560	.550	.525
50	3.78	1.19	.854	.722	.661	.622	.577	.537	.518	.500	.492	.469
60	3.45	1.09	.778	.658	.602	.567	.525	.490	.472	.456	.448	.428
80	2.98	.940	.672	.569	.520	.490	.454	.423	.408	.395	.388	.369
100	2.67	.840	.600	.508	.465	.438	.405	.378	.365	.353	.347	.329

When the computed value of k is smaller than the final entry in the column of the table, the sample size will be larger than 100 and may be computed as follows:

$$n = 100(K/k)^2, \tag{3}$$

where K is the tabular value for $n = 100$ in the column headed by m. In the illustrative example, if a were specified to be 0.6 instead of 1.5, the value of k would be 0.2 and n would be found to be

$$n = 100(.275/0.2)^2 = 189.$$

The example we have considered was stated so that a one-sided test of significance was implied. The experimenter was interested only in whether the mean exceeded the given figure, ten; he had no interest in a smaller mean however small it might have been. He would compute Student's t from his second sample and if it were greater than $t_{.10}(31) = 1.695$ he would conclude at the 5% level of significance that the mean exceeded ten. We have used the subscript .10 because the tables for critical levels of t ordinarily include both tails of the distribution in the critical region. The sample size 32 gives him 80% assurance that

$$t = \sqrt{32}\,(\bar{x} - 10)/s_2 \tag{4}$$

will exceed 1.695 if in fact the population mean is 11.5 ($\bar{x}$ is the sample mean and s_2 the sample standard deviation).

If one wishes to make two-tailed tests, then Tables I and II may be used provided the level of significance is doubled. Thus, to test whether the mean differs significantly in either direction from 10 in the illustrative example, one could again use the sample size 32 but the test would be at the 10% level of significance.

Now let us consider the possible errors that may be made in the conclusions in relation to β and γ which are .80 and .10 respectively in the problem at hand. If the population mean is ten, then the sample may give rise to a value of t larger than 1.695 with probability .05. The total probability of obtaining significance in either direction (the type I error of Neyman and Pearson) is .10. If the population mean is 8.5, the sample may give rise to a value of t which is not smaller than -1.695, and one will fail to detect at the .10 level of significance that the mean is smaller than ten; the probability of this error (the type II error of Neyman and Pearson) is $1-\beta=.20$ in the present instance. A similar situation occurs when the mean is actually 11.5; the probability is .20 that the sample will fail to show the mean to be significantly larger than ten at the .10 level of significance. When a type II error is made it will usually consist of the statement that the mean does not differ from ten. However there is a remote chance that an extreme error will be made—that the population mean will be declared smaller than ten when in fact it is 11.5, or declared larger than 10 when it is actually 8.5. The probability of such an extreme error is about .0006 for Table I and about .0002 for Table II. Thus in the given problem, if the true mean is 11.5 and one uses a sample of size 32 to test whether the mean differs from ten at the .10 level, the probability is about .0006 that the sample will give rise to a value of t which is smaller than -1.695, and thus lead one to the conclusion that the population mean is less than ten.

4. DIFFERENCES BETWEEN TWO MEANS

If one wishes to estimate the difference b between two population means knowing that the populations are normal and have the same variance, he may draw a sample of size $n+1$ from each population and put

$$t = \frac{(\bar{x} - \bar{x}' - b)\sqrt{n+1}}{\sqrt{s_2^2 + s_2'^2}} \tag{5}$$

equal to $\pm t_{1-\alpha}(2n)$ and solve for b to obtain the limits of the α confidence interval. ($\bar{x}$ and $\bar{x}'$ are the two sample means, and s_2^2 and s'^2_2 are the two sample variances.) Suppose one wants $n+1$ to be large enough that the probability will be β that the confidence interval will be shorter than $2d$, and that an estimate s_1^2 with m degrees of freedom has been previously obtained for one of the populations. The equation which determines n is

$$d^2/2s_1^2 = t^2_{1-\alpha}(2n)F_{1-\beta}(2n, m)/(n + 1) \tag{6}$$

as follows directly from equation (1). The divisor $2s_1^2$ arises from the fact that the variance of the difference between two observations is $2\sigma^2$ if each observation has variance σ^2. Equation (6) may be solved by trial and error for n.

In testing whether one mean is larger than another by an amount a one proceeds exactly as in the preceding section except that the quantity k to be found in the table is defined by

$$k = a/2s_1.$$

The number n thus determined is the number of degrees of freedom in the pooled estimate of the variance; the size of the sample to be drawn from each population is therefore $(n+2)/2$ if n is even or $(n+3)/2$ if n is odd. This sample size will insure that a difference in means of size a will be detected at the .05 level with probability .80 or .95 according to whether Table I or Table II is used to determine n. Here again it is assumed that the two populations are normal and have the same variance.

5 REPLICATION OF EXPERIMENTAL DESIGNS

Tables I and II may be used to aid one in making a decision about the number of replications to be used in a designed experiment. Suppose p normal populations with means $c_1, c_2, \cdots, c_p$ and a common variance are to be sampled; the means may represent treatment effects, for example, in an agricultural or industrial experiment. The most satisfactory way to deal with this problem is to have the experimenter specify what size treatment sum of squares he wishes to detect; that is, specify the size of

$$\sum_{i=1}^{p} (c_i - \bar{c})^2.$$

However, the solution of the problem in this form requires triple entry tables (the three variables being the numbers of degrees of freedom in the preliminary estimate of the error variance, the error variance of the experiment, and the treatment sum of squares). Since we had not the facilities to compute such tables, we shall have to consider the problem merely in terms of the differences between pairs of means.

Let us suppose that a number of treatments or treatment combinations are to be studied in an experimental design with r replications, that the number of degrees of freedom in the error variance s_2^2 for the proposed experiment is n, that any treatment mean will be based on q observations, and that an estimate s_1^2 with m degrees of freedom of the error variance is available. Two treatments with means $\bar{x}_i$ and $\bar{x}_j$ will be compared by computing

$$t = \sqrt{q}\,(\bar{x}_i - \bar{x}_j)/\sqrt{2}\,s_2$$

and comparing it with a critical level of t with n degrees of freedom. The numbers n and q will be functions of r.

The size of a difference between treatment means that has an 80% chance of being detected at the .05 level (in a given direction) may be computed for any proposed value of r by means of the relation,

$$\sqrt{\frac{2(n+1)}{}}\,ks_1, \tag{7}$$

where k is the entry in Table I corresponding to m and n. Similarly, using Table II to determine k, one may find the difference which has a 95% chance of being detected at the .05 level.

As an illustration, suppose six treatments are to be tested in a randomized block experiment with r replications and that we have an estimate, $s_1^2=9$, of the error variance with $m=8$ degrees of freedom. We then have $q=r$ and $n=5(r-1)$. For $r=2$, for example, a difference in treatment means of

$$a = \sqrt{6}\,(1.32)3 = 9.7$$

has an 80% chance of being detected in a particular comparison. With five replications, a is reduced to

$$a = \sqrt{\frac{42}{5}}\,(.624)3 = 5.4.$$

A few trial calculations of this nature for feasible values of r will provide one with an objective basis for coming to a decision about the number of replications to be used.

6. USE OF GUESSED ESTIMATES OF THE VARIANCE

It often happens that while an experimenter has no specific information about the variance of a distribution in the form of an estimate

s_1^2, he nevertheless has some more or less reliable opinion of its order of magnitude. He may, for example, have done considerable experimentation with several similar populations and thus have knowledge of the general range of variation in such populations. When an experimenter has information of this kind he would naturally like to use it, and of course he should do so—it will improve the efficiency of his experiments.

The question arises as to how one may quantify information of a general character; more precisely, how may one convert the information into numerical quantities which may be used for m and s_1^2? One simple method is for the experimenter to place what he feels to be reasonable upper and lower limits on the standard deviation. We may call these numbers G and H, and let us suppose that the experimenter has selected them in an effort to make it a fair nine-to-one bet that they will include the population standard deviation. In other words, G and H are to be regarded as 90% confidence limits for the standard deviation. One may then use the average of G and H for s_1:

$$s_1 = 1/2(G + H),$$

and determine the degrees of freedom, m, to be associated with s_1 by computing G/H and consulting Table III. The second line of this table consists merely of the square roots of the ratios of the .95 and .05 levels of chi-square.

TABLE III

DEGREES OF FREEDOM ASSOCIATED WITH VARIOUS RATIOS OF LIMITS ON THE STANDARD DEVIATION

m	1	2	3	4	5	6	7	8	9	10
G/H	31.3	7.6	4.7	3.7	3.1	2.8	2.6	2.4	2.3	2.2
g/h	3.6	2.2	1.84	1.67	1.58	1.51	1.46	1.42	1.39	1.36

If one has no firm feeling that his upper and lower limits really correspond to nine-to-one odds, he may prefer to select upper and lower limits, g and h, in such a way that he would have no choice between sides of an even-money bet that they include the standard deviation. Again the average of g and h would be employed as the estimate of s_1, and m would be determined by the third line of Table III which was computed using the .75 and .25 levels of chi-square. This method is, in fact, to be preferred over the first method because one may ordinarily be expected to have a more accurate conception of a fair even-money bet than a fair nine-to-one bet.

As an example, let us suppose 7 and 12 have been selected as 50% limits on a standard deviation. The estimate s_1 is 9.5, and since the ratio g/h is 1.7, there are about 4 degrees of freedom associated with the estimate. If one wishes to be conservative he may prefer to call it 3 degrees of freedom, but ordinarily the nearest figure may as well be selected. Only in rare occasions would there be any point in interpolating between 3 and 4, though of course there is no reason why a fractional value of m could not be used.

APPENDIX

A sample of size $n+1$ is to be drawn so that the probability will be β that an α confidence interval for the mean will be no longer than $2d$; an estimate s_1^2 of the variance with m degrees of freedom is available. If s_2^2 is the sample variance of the new sample, the half-length of the confidence interval will be $t_{1-\alpha}(n)s_2/\sqrt{n+1}$; it is desired that n be such that

$$P\left[t_{1-\alpha}(n)s_2/\sqrt{n+1} < d\right] = \beta \tag{8}$$

or

$$P\left[s_2 < \frac{\sqrt{n+1}\,d}{t_{1-\alpha}(n)}\right] = \beta. \tag{9}$$

For any n, it is known that

$$P(s_2 < \sqrt{F_{1-\beta}(n, m)}s_1) = \beta \tag{10}$$

where $F_{1-\beta}(n, m)$ is the critical level of the variance ratio for n and m degrees of freedom. On equating the right-hand sides of the inequalities in (9) and (10) one obtains the relation given in equation (1) for determining n.

We now turn to the derivation of Tables I and II. Suppose $\bar{x}$ is the sample mean and that the null hypothesis $H_0: \mu = 0$ is to be tested against alternatives $\mu > 0$ at the γ level of significance. H_0 would be rejected if

$$\sqrt{n+1}\,\bar{x}/s_2 > t_{2\gamma}(n) \tag{11}$$

where we have used the subscript 2γ in accordance with the usual notation of t tables. If the mean is actually $a > 0$ it is desired that (11) occur with probability β; we want

$$P\left[\sqrt{n+1}\,\bar{x}/s_2 > t_{2\gamma}(n)\right] = \beta$$

which may be written

$$P\left[t > t_{2\gamma}(n) - \sqrt{n+1}\,a/s_2\right] = \beta \tag{12}$$

where $t = \sqrt{n+1}(\bar{x} - a)/s_2$ has the Student distribution with n degrees of freedom. In (12) we shall put $D=\sqrt{n+1}\ a/s_1$ and $F=s_1^2/s_2^2$ to obtain

$$P[t > t_{2\gamma}(n) - D\sqrt{F}] = \beta. \tag{13}$$

We shall use this relation to determine a value D_0 for given values of m and n, then $D_0/\sqrt{n+1}$ corresponds to a particular value of a/s_1 given in advance, and $n+1$ will be the required sample size for that value of a/s_1. In determining D_0 we assume t and F have a distribution determined by the distributions of $\bar{x}$, s_1, s_2. This is legitimate for only one use of s_1; clearly for repeated use of s_1 to determine sample sizes one would have to evaluate D from (13) using the conditional distribution of F given s_1, and that conditional distribution involves the unknown variance. We thus in effect treat the problem as if the question of determining the same size arose in advance of the information in s_1. (An excellent discussion of this general problem has been given by Stein in reference 4.)

The variates t and F in (13) are not independent since they are both functions of s_2; however their joint distribution is readily found as follows: Since $\bar{x}$, s_1, and s_2, are mutually independent their joint density function is the product of their individual densities. On transforming the three variates to t, F, and s_2, and then integrating out s_2 one finds the density of t and F to be

$$f(t, F) = CF^{(m-2)/2}/(t^2 + mF + n)^{(m+n+1)/2} \tag{14}$$

C being a function of m and n.

To facilitate the numerical integration needed to evaluate the integral implied by (13), further transformations were made. If one lets

$$x^2 = (m + n)t^2/(mF + n)$$
$$y = mF/(mF + n)$$

the density function reduces to

$$g(x, y) = \left[\frac{1}{\sqrt{\pi(m+n)}} \frac{\left(\Gamma\frac{m+n+1}{2}\right)}{\Gamma\left(\frac{m+n}{2}\right)\left(1+\frac{x^2}{m+n}\right)^{(m+n+1)/2}}\right] \left[\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m}{2}\right)} \cdot y^{(m-2)/2}(1-y)^{(m-2)/2}\right.$$

so that x and y are independently distributed and by tabulated functions, x having the Student distribution and y the beta distribution. Nevertheless one numerical integration is required to evaluate the probability in (13), and since we fixed β in advance it was necessary to do several such integrations for trial values of D to determine each entry in Tables I and II. $D/\sqrt{n+1}$ rather than D is tabulated. The entries corresponding to $m=1, 2, 4, 8, 16, 32$ and $n=1, 2, 4, 16, 32, 64, 128$ were determined by integration; the other entries were filled in by interpolation. The level of accuracy of the tabular values is about one in 300; an entry such as 9.12 may be in error by three units in the last place. The entries in the right hand column of each table were taken from the Neyman and Tokarska (reference 3) table of the power of the t test.

REFERENCES

[1] R. A. Fisher and F. Yates, *Statistical Tables*, Oliver and Boyd, London, 1938.
[2] Maxine Merrington and Catherine M. Thompson, "Percentage Points of the Inverted Beta Distribution," *Biometrika*, Vol. 33, 1943, p. 73.
[3] J. Neyman and B. Tokarska, "Errors of the Second Kind in Testing Student's Hypothesis," *Journal of the American Statistical Association*, Vol. 31, 1936, p. 318.
[4] Charles Stein, "A Two-Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance," *Annals of Mathematical Statistics*, Vol. 16, 1945, p. 243.

APPLICATION OF THE THEORY OF EXTREME VALUES IN FRACTURE PROBLEMS*

BENJAMIN EPSTEIN
Coal Research Laboratory, Carnegie Institute of Technology

In this paper it is shown that the theory of extreme values is pertinent to the treatment of certain aspects of the fracture or break-down of materials used in modern technology. An attempt is made to integrate some of the results scattered through the technical literature.

INTRODUCTION

DURING THE PAST twenty-five years there has been an increasing interest in gaining a deeper insight into the laws governing some of the phenomena observed in connection with the fracturing of metals, textiles, glass and other materials used in modern technology. Among other things, one would like to account for such things as the skewness of the distribution of breaking strengths of supposedly identical samples and the dependence of strength on specimen size.

It appears, from a survey of the literature on this subject, that workers in many different fields have obtained special results without realizing fully the general setting of the problem. It is our purpose in this paper to indicate how the application of the theory of extreme values leads to a resolution of some of the pertinent questions.

SURVEY OF FIELD

In essence, all the statistical models proposed in the study of fracture take as a starting point Griffith's theory[1] which states that the difference between the calculated strengths of materials and those actually observed resides in the fact that there exist flaws in the body which weaken it. Accepting this point of view is equivalent to assuming that there will be a distribution of strengths in a given specimen in the sense that a different amount of force will be needed to fracture a specimen at one or another point. There are many physical situations where failure at one point means failure of the entire specimen. But this simply means statistically that it is the worst flaw among N flaws (where N is the number of flaws in the specimen) which determines the strength of a specimen and therefore the theory of extreme values is immediately applicable. Clearly N increases as the specimen size increases and, therefore, the dependence of strength on specimen size is equivalent statistically to the problem of how the distribution of smallest (or largest) values depends on N. Generally speaking, asymptotic theory is applicable since N, the number of flaws, is large in most practical problems.

---

* Presented at the 107th Annual Meeting of the American Statistical Association, December 30 1947, in connection with the program on "The Theory of Extreme Values and Its Applications."

[1] A A Griffith, "The Phenomena of Rupture and Flow in Solids," *Philosophical Transactions of the Royal Society, 221A*, 163 (1920), "The Theory of Rupture," *Proceedings of the First International Congress for Applied Mechanics, 1*, Delft, p 55 (1924)

As far as is known to the author, the first worker to realize the connection between specimen strength and distribution of extreme values was F. T. Peirce[2] of the British Cotton Industry Association. The application of essentially the same ideas to the study of the strength of materials was carried out by the Swedish physicist and engineer, Weibull.[3] He, unlike Peirce who assumed a Gaussian distribution of strengths in the locality of flaws, made the assumption that $F_0(S)$, the probability of breakage as a function of the stress S is given by $F_0(S)=1-e^{-(S/S_0)^m}$, where S_0 and m are unknown parameters depending on the characteristics of the material under test. Experimental data drawn from the fields of textiles, metallurgy, dielectrics, and metal fatigue indicate that Weibull's assumptions are reasonable. Recently his views were partially corroborated by Davidenkow, Shevandin and Wittman.[4]

The Russian physicists, Frenkel and Kontorova,[5] were the next to study these problems. Starting with a distribution of flaw strengths x given by the probability density function,

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma}$$

they found the most probable value of the strength of specimens of volume V large enough to contain many flaws is approximately given by

$$(1) \quad \mu - \sigma\sqrt{2 \log nV - 2 \log 2\sqrt{\pi}}$$

where n is the average number of flaws per cubic centimeter of the material. This result states that the strength of specimens under the Frenkel-Kontorova assumptions should decrease linearly with $\sqrt{\log V}$.

---

[2] F. T. Peirce, "Tensile Tests for Cotton Yarns V. 'The Weakest Link'—Theorems on the Strength of Long and of Composite Specimens," *Journal of the Textile Institute*, Transactions, *17*, 355 (1926).

[3] W. Weibull, "A Statistical Theory of the Strength of Materials," *Ing. Vetenskaps Akad. Handl.*, No. 151 (1939); "The Phenomenon of Rupture in Solids," *Ibid.*, No. 153 (1939). See also John Tucker, "Statistical Theory of the Effect of the Dimensions and Method of Loading upon the Modulus of Rupture of Beams," *Proceedings of the American Society of Testing Materials*, *41*, 1072 (1941).

[4] N. Davidenkow, E. Shevandin and F. Wittman, "The Influence of Size on the Brittle Strength of Steel," *Journal of Applied Mechanics*, *14*, No. 1, A63–67 (1947).

[5] T. A. Kontorova, *J. Tech. Phys. U.S.S.R.*, *10*, 886 (1940); J. I. Frenkel and T. A. Kontorova, "A Statistical Theory of the Brittle Strength of Real Crystals," *J. Phys. U.S.S.R.*, *7*, 108 (1943).

This should be compared with Weibull's result to the effect that the most probable value of the strength of specimens of volume V is proportional to $S_0V^{-1/m}$. Frenkel and Kontorova consider the discovery of (1) to be very important and to justify the assumption that the local distribution of strengths due to flaws is Gaussian. Statisticians familiar with the theory of extreme values will recognize (1) as due in more precise form to R. A. Fisher and Tippett[6] and Gumbel.[7] The more precise formula for (1) is

$$\mu - \sigma\sqrt{2 \log \mathrm{n}V} + \sigma \frac{\log \log nV + 4\pi}{2\sqrt{2 \log nV}} .$$

Beginning with World War II several scientists in this country became interested in statistical theories of fracture. Some work was done by Ruark and Rosen[8] of the University of North Carolina under contract with the Navy Department. More recently two metallurgists, J. C. Fisher and Hollomon,[9] have proposed a statistical theory of fracture based on the idea that the material under study is an elastic solid containing many thin disc-like cracks with elliptical cross-sections. The major axes of the ellipses are assumed to be of size x possessing distribution $p(x) = he^{-hx}$, $x \geqq 0$. Brooks[10] of the Westinghouse Electric Corporation has presented experimental evidence indicating the importance of the theory of extreme values in understanding the breakdown of paper capacitors. Collaboration between Brooks and the author[11] has resulted in a paper which contains results useful to design engineers in the field of dielectrics. It is shown, among other things, that the most probable value of the breakdown strength $\tilde{V}(A)$ of capacitors of area A is related to A by a relationship of the form

$$(2) \qquad \tilde{V}(A) = \alpha - \beta \log A,$$

where α and β are positive constants determined by the number of flaws (technically known as conducting particles) per unit area and by the particle size distribution.

---

[6] R. A. Fisher and L. H. C. Tippett, "Limiting Form of the Frequency Distribution of the Largest and Smallest Member of a Sample," *Proceedings of the Cambridge Philosophical Society*, *24*, 180 (1928).

[7] E. J. Gumbel, "Les valeurs extrêmes des distributions statistiques," *Annales de l'Institut Henri Poincaré*, *5*, 115 (1936).

[8] A. E. Ruark and N. Rosen, "Statistical Effects in the Testing of Materials," *Naval Research Laboratory Report No. 0-2191*.

[9] J. C. Fisher and J. H. Hollomon, "A Statistical Theory of Fracture," American Institute of Mining and Metallurgical Engineers, Tech. Pub. No. 2218, *Metals Technology*, August 1947.

[10] H. Brooks, "The Probable Breakdown Voltage of Paper Dielectric Capacitors," *American Institute of Electrical Engineers, Technical Paper*, 47–164, 1947.

[11] B. Epstein and H. Brooks, "The Theory of Extreme Values and Its Implications in the Study of the Dielectric Strength of Paper Capacitors," accepted by the *Journal of Applied Physics*.

All of the papers mentioned thus far were based essentially on the concept that fracture is determined by the weakest link. A much more complicated and very interesting problem has recently been considered by H. E. Daniels.[12] This paper is devoted to studying the strength of bundles of threads where the load S is now applied to a set of elements in parallel and not in series. Whereas in the series case each element carries the full load S, in the parallel case each thread initially carries load S/n. It turns out that a bundle will not break under a load S if, and only if, there exists an integer $k \leqq n$ such that k among the threads have strength exceeding S/k. The problem of the distribution of strength of bundles of n threads can no longer be treated as a problem in the distribution of smallest values. It turns out that whereas the distribution of strengths of single long strands is negatively skewed, the distribution of strengths of bundles of n threads is asymptotically normal for large n under fairly general assumptions about the distribution of strengths of individual strands.

Daniels' result has been mentioned here even though it is not a purely "weakest link" problem because of its possible importance in studying certain aspects of the strength of materials. In any real problem the situation is probably neither precisely described by elements in series or parallel, but rather by elements distributed in some rather complicated way. However, it is of value to know what to expect in certain limiting idealized models.

THE IMPLICATIONS OF VARIOUS DISTRIBUTION ASSUMPTIONS IN BREAKDOWN PROBLEMS

There is considerable dispute in the technical literature as to which assumptions concerning the distribution of flaw strengths or distributions of sizes of cracks or conducting particles are the correct ones. Of course, this sort of question cannot be answered purely on statistical grounds. Adequate experimental data of the right type must be available before one can decide, in a reasonably rigorous way, what the underlying distribution laws are. At the moment the situation seems to be as follows: (a) there is a certain amount of data drawn from a variety of fields which partially substantiates Weibull's assumptions; (b) many experimenters have observed that the most probable value of the strength decreases as some function of the logarithm of the size of specimen; (c) the distribution of strengths of specimens all of the same size appears to be negatively skewed; (d) in the breakdown of capacitors there appears to be ample justification for assuming that the sizes of conducting particles are distributed according to a size distribution of the form $f(x)=\lambda e^{-\lambda x}$, $x>0$. It can then easily be shown that the most probable value of the breakdown voltage will decrease linearly with the logarithm of the area. Physical data substantiate this and other conclusions.

---

[12] H. E. Daniels, "The Statistical Theory of the Strength of Bundles of Threads. I," *Proceedings of the Royal Society*, London, *183*, 405, 1945.

The physical implications of various distribution assumptions are discussed in a recent paper.[13] In treating such questions one is generally justified in working with the asymptotic portion of extreme value theory since the number of defects involved in fracture problems is large. It is a very important fact that the distribution of extreme values in samples of size n ($n\to\infty$) possesses a kind of stability with n playing in a certain sense the role of a parameter. It is shown, for example, in Cramér[14] that if samples of size n are drawn from a population described by the continuous p.d.f. $f(x)$ and associated d.f. $F(x)=\int_{-\infty}^{x} f(x)\,dx$, then the asymptotic frequency distribution of y, the smallest value in samples of size $n(n\to\infty)$ is given by

$$h(y) = nf(y)e^{-nF(y)}. \tag{3}$$

The d.f. is

$$H(y) = 1 - e^{-nF(y)}. \tag{4}$$

It is convenient to introduce a random variable η which is related to the random variable y by the relationship $\eta=nF(y)$, $0\leqq\eta\leqq n$. (3) then assumes the simple form

$$h(\eta) = e^{-\eta}, \qquad 0 \leqq \eta \leqq n. \tag{5}$$

Similarly, the asymptotic p.d.f. of x, the largest value in sample of size n, is given by

$$h(x) = nf(x)e^{-n[1-F(x)]}. \tag{6}$$

The asymptotic d.f. is

$$H(x) = e^{-n[1-F(x)]}. \tag{7}$$

If a random variable $\xi=n[1-F(x)]$ is introduced, then (6) becomes

$$h(\xi) = e^{-\xi}, \qquad 0 \leqq \xi \leqq n. \tag{8}$$

Equations (3) through (8) form a basis from which a number of physical conclusions of interest can be drawn. Equation (4) has a direct physical interpretation in the class of problems under consideration. Imagine that one is breaking one-dimensional threads possessing Δ flaws per unit length and with flaw strength distribution $f(x)$. Then ΔL is the expected number of flaws in a specimen of length L and $\Delta LF(x)$ is the expected number of flaws having strength $\leqq x$ in specimens of length L. But, as is presumably (but not necessarily) the case, the flaws are distributed individually and collectively at random, and therefore it can be assumed that the number of flaws having strength $\leqq x$ in specimens of length L is a random variable distributed according to the law of Poisson. Therefore, the expected fraction of specimens of length L having strength $\leqq x$ is given *exactly* and not merely asymptotically by

---

[13] B. Epstein, "Statistical Aspects of Fracture Problems," *Journal of Applied Physics, 19*, 140–147 (1948).

[14] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, pp. 373–378, 1946.

$$(9) \qquad H(x) = 1 - e^{-\Delta LF(x)}.$$

The identification of (9) with (4) is complete since ΔL plays the role of the sample size n. A similar interpretation can be attached to (7).

If $f(x)$ is assumed to be one of a number of common distribution functions, the application of (4) and (5) yields the results in Table I. The results for the first two distributions will be found in Cramér.[14]

TABLE I

Probability Density Functions	Asymptotic Distribution of Smallest Values in Samples of Size n (n large)
1. Laplace's distribution $f(x) = \dfrac{1}{2\lambda} e^{-\lvert x-\mu\rvert/\lambda}$	$y = \mu - \lambda \log n/2 + \lambda \log \eta$ or $y = \mu - \lambda \log n/2 + \lambda\zeta$
2. Gauss' distribution $f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$	$y = \mu - \sigma\sqrt{2 \log n}$ $+ \sigma \dfrac{\log \log n + \log 4\pi}{2\sqrt{2 \log n}} + \dfrac{\sigma \log \eta}{\sqrt{2 \log n}}$ or $y = \mu - \sigma\sqrt{2 \log n}$ $+ \sigma \dfrac{\log \log n + 4\pi}{2\sqrt{2 \log n}} + \dfrac{\sigma\zeta}{\sqrt{2 \log n}}$
3. Distribution of type assumed by Weibull $f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, x \geqq 0, \alpha > 0, \beta > 1$	$y = (\eta/n\alpha)^{1/\beta}$ or $\log y = -1/\beta \log n\alpha + \zeta/\beta.$

η is distributed with p.d.f. $h(\eta) = e^{-\eta}$, $\eta \geqq 0$ and ζ is distributed with p.d.f. $g(\zeta) = e^{\zeta} e^{-e^{\zeta}}$, $-\infty < \zeta < \infty$

Since ζ, the random portion of y, the smallest value in samples of size n, is distributed as $e^{\zeta} e^{-e^{\zeta}}$, a distribution which is strongly skewed to the left, it can be said that: (a) if the distribution of flaw strengths follows the Laplace law, Gaussian law, or more generally one of the type $Ae^{-B\,|x-\mu|^p}$ for large values of $|x-\mu|$ (where A, B, μ, and p are positive constants), then the distribution of strengths of samples all of the same size is negatively skewed. More precisely, if $P(y)$ is the fraction of specimens of strength $\leqq y$, then log log $1/1-P(y)$ is a linear function of y; (b) if the distribution of flaw strengths follows a Weibull distribution then the distribution of the logarithms of the strength of samples all of the same size is negatively skewed. More precisely, if $P(y)$ is the fraction of specimens of strength $\leqq y$, then log log $1/1-P(y)$ is a linear function of log y.

There are physical data at present available which follow each of these patterns. For example, the distribution of breakdown voltages of capacitors[10,11] all of the same size is of type (a). Weibull[8] gives several examples, drawn from a number of fields, which are of type (b).

In the study of size effect, i.e., how specimen size affects the distribution of strengths, it is important to see what happens in Table I when the size of sample changes. In particular, it is of interest to see how the most probable value (i.e., the mode), the mean, and variance of the smallest value in samples of size n change with n. To do this note that

$$\text{(10) The most probable value of } \zeta = \tilde{\zeta} = 0,$$

$$\text{The mean value of } \zeta = E(\zeta) = \int_0^\infty (\log z)e^{-z}dz$$

$$= -.577 \text{ and}$$

$$\text{The variance of } \zeta = E(\zeta^2) - [E(\zeta)]^2$$

$$= \int_0^\infty (\log z)^2 e^{-z}dz$$

$$- \left[\int_0^\infty (\log z)e^{-z}dz\right]^2$$

$$= \pi^2/6,$$

and

$$\text{(11) The most probable value of } \eta^{1/\beta} \text{ is } [(\beta-1)/\beta]^{1/\beta},$$

$$E(\eta^{1/\beta}) = \Gamma\left(\frac{\beta+1}{\beta}\right) \text{ and}$$

$$\text{var. } (\eta^{1/\beta}) = \Gamma\left(\frac{\beta+2}{\beta}\right) - \left[\Gamma\left(\frac{\beta+1}{\beta}\right)\right]^2.$$

It then follows that

$$\text{(12)} \quad \tilde{y} = \mu - \lambda \log(n/2), \qquad \bar{y} = \tilde{y} - 0.577\lambda,$$

$$\sigma_y{}^2 = \lambda^2\pi^2/6,$$

if the original population is a Laplace distribution with parameters (μ, λ),

$$(13)\quad \tilde{y} = \mu - \sigma\sqrt{2\log n} + \sigma\,\frac{\log\log n + 4\pi}{2\sqrt{2\log n}},$$

$$\bar{y} = \tilde{y} - 0.577\sigma/\sqrt{2\log n},$$

$$\sigma_y^2 = \sigma^2\pi^2/12\log n,$$

if the original population is Gaussian with parameters (μ, σ), and

$$(14)\quad \tilde{y} = \frac{1}{(\alpha n)^{1/\beta}}\left(\frac{\beta-1}{\beta}\right)^{1/\beta} = \tilde{x}n^{-1/\beta},$$

$$= \Gamma\left(\frac{\beta+1}{\beta}\right)\Big/(n\alpha)^{1/\beta},$$

$$\sigma_y^2 = \left\{\Gamma\left(\frac{\beta+2}{\beta}\right) - \Gamma^2\left(\frac{\beta+1}{\beta}\right)\right\}\Big/(n\alpha)^{2/\beta},$$

if the original population is of the Weibull type with parameters (α, β). $\tilde{x}$ denotes the mode of the original Weibull distribution.

Equations (12), (13) and (14), when interpreted physically, give definite information on how the size effect depends on the distribution of strengths in the vicinity of flaws. In each instance the specimens become weaker as V, the specimen size, increases. In the first two cases the relationships are both semi-logarithmic in character ($\tilde{y}$ decreases linearly with $\log V$ in the Laplace case and with $\sqrt{\log V}$ in the Gauss case). However, the spread of the distribution remains unchanged in the Laplace case, but becomes narrower with increasing V in the Gauss case. Under the Weibull assumption $\log \tilde{y}$ (as well as $\log y$) decreases linearly with $\log V$ and the spread also decreases.

It can be shown similarly that if the flaw strength distribution is of the form $Ae^{-B|x-\mu|^p}$ for large values of $|x-\mu|$ (where A, B, μ, and p are positive constants), then the most probable value of the strength will decrease proportionately to $(\log n)^{1/p}$. If the distribution is rectangular then there would be no size effect. The results in this paragraph, in Table I, and in (12), (13), and (14) furnish a statistical basis for testing which assumptions on the underlying flaw strength distributions are reasonable. In the last analysis definitive answers of physical interest must await the accumulation of more pertinent physical data. Some of the techniques used in analyzing such data are closely related to those used by Gumbel on the theory of floods[15] and distribution of oldest ages at death.[16]

It should be mentioned that in some strength problems it is convenient to work with the distribution of largest values in preference to the distribution of smallest values. This is particularly true in studying the breakdown strength of capacitors. It is fairly well established that one of the major causes of breakdown of capacitors is the presence of flaws known technically as conducting particles. These particles are present no matter how carefully capacitor paper is made or how carefully the capacitor itself is manufactured. The conducting particles are spread at random throughout the area of the capacitor and, depending on their size and orientation, create a local weakening of the capacitor by reducing the nominal insulation thickness in the neighborhood of the flaws. But the voltage required to break down a capacitor is equal to the voltage required to break it down at its weakest point, or equivalently at the place where the greatest effective penetration of the insulation has occurred. Therefore, the distribution of largest conducting particle sizes is pertinent in this problem. Working with largest particle sizes in a sample also appears to be a good way to treat the metallurgical problem of Fisher and Hollomon.[9] If this is done it turns out that their statistical and physical assumptions taken together imply that the strength of specimens of size V should fall off as the reciprocal of $\sqrt{\log V}$.

CONCLUSIONS AND GENERAL REMARKS

When the problem of fracture is viewed in its most general setting it is clear that the "weakest link" concept, when applicable, may be the key to the occurrence of certain observed relationships between the strength of specimens and their size. The term strength is meant to be quite general in scope: we may have in mind mechanical strength, electrical strength, resistance of painted specimens to the corrosive effects of the atmosphere, or perhaps the life span of a device which ceases to function when any one of a number of vital parts breaks down. From a statistical point of view phenomena of this type, though different at first glance, are really equivalent since they lead in each instance to the same problem, namely, the distribution of smallest values in large samples. This may account for the fact that workers in many different fields have observed skewed strength distributions and approximately semi-logarithmic dependence on size. As a matter of fact, the consider-

---

[15] E. J. Gumbel, "The Return Period of Flood Flows," *Annals of Math Stat.*, *12*, 163, 1941.

[16] E. J. Gumbel, "La durée extrême de la vie humaine," *Actualités Scientifiques et Industrielles*, Hermann et Cie, Paris, 1937.

ations in this paper indicate that under certain hypotheses there is a *statistical compulsion* for such results.

Many scientists accept the fact that the Gaussian distribution plays a fundamental role in science, and, in fact, there are many who feel that this is the only distribution which nature calls truly her own. It has been shown in this paper, however, that in a certain class of phenomena the typical distributions are far from normal, and are, in fact, often strongly skewed to the left. There are other instances lying outside the scope of the present considerations where still other "characteristic" distributions are found. For example, the particle sizes arising from crushing and grinding have a marked tendency to follow a logarithmico-normal form.

In conclusion we wish to emphasize that the "weakest link" hypothesis cannot be expected to explain all phenomena related to strength of materials. In this paper it has been assumed that the flaws are distributed at random and do not influence one another in any way. Either of these assumptions may be totally or in part untrue in certain physical situations. For example, putting a notch on the surface of a beam subjected to bending may cause the strength to drop a great deal. However, if many notches are put on the surface, there is scarcely any effect on strength. This fact is of importance in the fatigue strength of a body having a highly polished surface. A single scratch may be ruinous whereas roughing up the whole surface with many scratches may bring about little change in the strength.

Weibull[3] and Daniels[12] have obtained some interesting results relating to more general situations but much remains to be done.

PROBLEMS WITH SAMPLING PROCEDURES FOR RESERVE VALUATIONS*

GEORGE C. CAMPBELL
Metropolitan Life Insurance Company

The computation of the reserve liability, running up to about $49 billion in North American life insurance companies, involves a large amount of work, some of which might be saved if sampling methods were satisfactory. Certain possible approaches to the sampling problem are mentioned, and the mathematical problem is developed for one such approach. Experimental results are presented from one small exploratory sample intended to establish orders of magnitude. Certain tentative conclusions are presented.

It is well known that the death rate is very low in the teens; that it is fairly low but increasing slowly up to the middle ages; and that it then increases rapidly through the very old ages. Theoretically, it is possible to increase the premium for life insurance each year of age in accordance with the risk of death in that year, starting with a very low premium in the beginning and reaching very high annual premiums late in life. Such plans have been tried and found impractical. Large and rapidly increasing premiums late in life are unrealistic in relation to the policyholder's income.

Very early in the insurance business it was found more practical to sell insurance on a level annual premium basis during the premium-paying period. This means that the level annual premiums are more than necessary to pay expected claims in the early years of the insurance. These amounts in excess of claims must be saved and invested to provide a fund for the later years of insurance when the expected claims will exceed the premium collections. The theory is relatively simple for the computation of this fund which is called the reserve. When a policy is issued, the present value of future net premiums, after deducting the loading for expenses in the gross premium, is exactly equal to the present value of the benefit to be paid. In computing present value, future payments are discounted to the present time on the basis of interest and the probability of survival according to a mortality table. Considering the case where the risk of death is increasing with age, if the present value of the level benefit is exactly equal to the present value of the net level premiums when a policy is issued, then at any time subsequent to issue, the present value of the benefit will exceed the present value of the future net premiums. The excess of the present value of the benefit over the present value of future net premiums at any point in time is the reserve on the policy which the Company must set up in order to meet its commitments.[1] The reserve is a definite liability of the Company which must be shown on the balance sheet as required by law.

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

After the mortality table and the rate of interest have been selected to satisfy legal and managerial requirements, the average net level premium reserve per $1 of insurance is uniquely determined for each type of policy by the contractual provisions of the policy (including statutory references), by the duration from issue, and by the age at issue. In the Metropolitan, such a classification leads to about 200,000 individual cells for Ordinary and Industrial Life Insurance. It is necessary to ascertain the amount of insurance in force (i.e., in being) falling into each of these cells. A large amount of work is expended in the classification of all current transactions, such as new policies issued, lapses, revivals, death claims, maturities, etc., into these 200,000 cells in order to carry forward the classification from one year to the next. Even after the insurance in force is obtained in this detail, it is necessary then to multiply the amount of insurance in each one of these cells by the appropriate reserve factor per dollar of insurance to obtain the corresponding reserve liabilities which in turn must be summarized into final figures. All of this work must be thoroughly checked in detail and by controls to insure accuracy.

The determination of the reserve liability for all of the policies the Company has outstanding is generally referred to as reserve valuation. Separate calculations are made for the several branches of business, and for related benefits. The amounts involved are very large and constitute the principal liability of any life insurance company. For the business at large, in North America at the end of 1947, the total assets amount to about $55 billion of which about $49 billion is required to cover reserve liabilities. In any one of the larger companies, the reserve may well run to a few billion dollars. It is clear that the reserve valuation problem is of great practical importance and that it involves considerable labor and expense. It should not be assumed, however, that most of this work could be eliminated simply by eliminating or modifying the requirements for reserve valuation. Much of the same information would still be needed for mortality studies and for other purposes. If it were possible to obtain the reserve liability from a reasonably small sample with sufficient accuracy to protect adequately the policyholders and to be accepted by Company management and by State supervisory authorities, a considerable saving in Home Office expenses could be made.

---

[1] For details regarding net premium and reserve computations, see E. F. Spurgeon, *Life Contingencies*, Cambridge University Press, 1932.

THE MATHEMATICAL PROBLEM OF SAMPLING

When we approach the valuation problem from the sampling point of view, a number of possibilities present themselves. The maximum saving in clerical labor could be made if a satisfactory approximation could be obtained by sampling from a file of cards representing policies in force without the need for recourse to any running transaction accounts whatsoever. This would mean, for a particular block of business, that we would require reliable estimates from a sample alone for a number of policies, the amount of insurance, and the reserve, together with their sampling variances.

It seems unlikely, however, that Company management would be willing to discard all control accounts on the number of policies and on the amount of insurance. Consequently, it seems more realistic to approach the sampling problem on the assumption that certain control accounts would be kept, giving the number of policies and the amount of insurance in force accurately for certain blocks of business in convenient sub-groups. The problem reduces to finding the average reserve per $1 insurance, and its variance, from a sample separated into corresponding sub-groups. The reserve would then be estimated for each sub-group from the actual total amount of insurance and the average reserve per $1 insurance as estimated from the sample.

It appears desirable to take advantage of the high correlation existing between the amount of reserve and the amount of insurance on groups of individual policies. There is an exact mathematical relationship for individual policies, and a correlation coefficient of the order of 0.7 for the aggregate group of Regular Ordinary policies in the Metropolitan. In most of the sampling problems related to valuation studied by the writer, the solution depends on the relation of the quantity being estimated (in this case, the reserve) either to the amount of insurance or to the amount of gross annual premium. This relationship leads to a two-way scatter diagram.

Consider the two-way scatter diagram presented as Table I with the reserve factors per $1 of insurance along the X-axis and the amount of insurance along the Y-axis. In the mathematical derivation we must consider the related scatter diagram formed by the amount of reserve along the X-axis and the amount of insurance along the Y-axis. Fortunately, numerical results for the second scatter diagram can be obtained from the first scatter diagram based on reserve factors by taking the appropriate cross product between the reserve factor and the amount of insurance whenever the amount of reserve is needed.

TABLE I

STRIBUTION OF A SAMPLE OF 1,015 REGULAR ORDINARY POLICIES ACCORDING TO AMOUNT OF INSURANCE AND RESERVE PER DOLLAR OF INSURANCE

Reserve per Dollar of Insurance (*X*-Axis)	unt of Insurance per Policy, Dollars (*Y*-Ax														
	500	1,000	1,001 −1,999	2,000	2,001 −2,999	3,000	3,001 −3,999	4,000 −4,999	5,000	5,001 −5,999	6,000 −6,999	7,000 −9,999	10,000	10,001 −24,999	25,000	Totals
.000–.024		8	1	1	8	0	1	1	15		2	1	6	1		45
.025–.049		46	4	8	0	0	1	2	4	1			2			68
.050–.074		61	7	6	6	2			6				2	1		91
.075–.099		61	6	6	6	0	1		11		1	1	1			94
.100–.124		66	11	10	5	2	1		7	1	1		1			105
.125–.149		46	3	5	4	1	1		7	1						68
.150–.174		35	5	7	3	1	1	2	12	2	1					69
.175–.199		30	3	4	2	1			3				3			46
.200–.224		26	3	5	3	0			3				2	1	1	44
.225–.249		24	4	4	2	3		1	6	1						45
.250–.274		19	5	6	2	0			7		1	1				41
.275–.299		24	3	1	0	1			1	1				1		32
.300–.324		15	0	4	2	0			6				1			28
.325–.349		13	1	5		0			5	1			1		1	27
.350–.374		13	1	5	1	2			3		1					26
.375–.399		15	1	6		1	1		1							25
.400–.424		11	4	5	2				5				1			28
.425–.449		9	1	3					1							14
.450–.474	1	13	1	2			1		1				1			20
.475–.499		5	1	3					2							11
.500–.524		8	1	1		1	1		2			1			1	16
.525–.549		7	1	2					2							12
.550–.574		3	1	4	1				2							11
.575–.599		2	2	0					0							4

TABLE I—*Continued*

DISTRIBUTION OF A SAMPLE OF 1,015 REGULAR ORDINARY POLICIES ACCORDING TO AMOUNT OF INSURANCE AND RESERVE PER DOLLAR OF INSURANCE

P

Reserve per Dollar of Insurance (*X*-Axis)	500	1,000	1,001 -1,999	2,000	2,001 -2,999	3,000	3,001 -3,999	4,000 -4,999	5,000	5,001 -5,999	6,000 -6,999	7,000 -9,999	10,000	10,001 -24,999	25,000	Totals
.600–.624		2	1	2		1			2							8
.625–.649		1		0					1							2
.650–.674		3		1					1							5
.675–.699		0			1				0							1
.700–.724		2							1							3
.725–.749		3							0							3
.750–.774		1							0							1
.775–.799		2		1					1							4
.800–.824		0														0
.825–.849		0														0
.850–.874		3														3
.875–.899		1														1
.900–.924		4		2												6
.925–.949		0														0
.950–.974		0														0
.975–.999		5	1	1												7
1.223		1														1
Totals	1	588	72	110	48	16	9	6	118	8	7	4	21	4	3	1,015

The average reserve per $1 insurance in a sub-group of any particular block of business is obtained readily by summing the reserve on all of the policies in the sub-group and dividing by the sum of the insurance on all of its policies. An estimate is required for the variance of the average reserve per $1 of insurance.

Let $\bar{V}$ = the average reserve per $1 insurance
$\bar{R}$ = the average reserve per policy
$\bar{S}$ = the average insurance per policy

Then

$$\bar{V} = \bar{R}/\bar{S} \text{ for the universe} \tag{1}$$

and

$$\bar{V} + \Delta\bar{V} = \frac{\bar{R} + \Delta\bar{R}}{\bar{S} + \Delta\bar{S}} \text{ for a sample} \tag{2}$$

Hence

$$\Delta\bar{V} = \frac{\bar{R}}{\bar{S}}\left[\frac{\Delta\bar{R}}{\bar{R}} - \frac{\Delta\bar{S}}{\bar{S}}\right]\left[1 + \frac{\Delta\bar{S}}{\bar{S}}\right]^{-1} \tag{3}$$

and

$$(\Delta\bar{V})^2 = \frac{\bar{R}^2}{\bar{S}^2}\left[\frac{(\Delta\bar{R})^2}{\bar{R}^2} - 2\frac{\Delta\bar{R}\Delta\bar{S}}{\overline{RS}} + \frac{(\Delta\bar{S})^2}{\bar{S}^2}\right]\left[1 + \frac{\Delta\bar{S}}{\bar{S}}\right]^{-2}. \tag{4}$$

Summing for all samples, dividing by the number of samples, and taking unity as a reasonable approximation to the last factor

$$r_{\bar{V}}^2 = \frac{\bar{R}^2}{\bar{S}^2}\left[\frac{\sigma_{\bar{R}}^2}{\bar{R}^2} - 2\frac{\sigma_{R}\sigma_{\bar{S}}r}{\overline{RS}} + \frac{\sigma_{S}^2}{\bar{S}^2}\right]. \tag{5}$$

Replacing the variances of averages by the variances of individual policies, we have the approximate formula

$$\frac{\bar{R}^2}{n\bar{S}^2}\left[\frac{\sigma_{R}^2}{\bar{R}^2} - 2\frac{\sigma_{R}\sigma_{S}r}{\overline{RS}} + \frac{\sigma_{S}^2}{\bar{S}^2}\right] \tag{6}$$

where n is the number of policies in the sample and r is the correlation coefficient between R and S.

The above expression enables us to compute an approximate value of the required variance of the average reserve per $1 of insurance from the scatter diagram.

EXPERIMENTAL WORK

Some exploratory work was done with a sample of 1,015 Regular Ordinary policies drawn from a total in force of about 4.3 million policies for $9.6 billion of insurance with $2.0 billion of reserve liability. It was thought that a small exploratory sample would serve as a pilot study to establish certain orders of magnitude which might be helpful in planning any subsequent investigation of the sampling problem. There is no intention to use such a small sample for actual valuation purposes.

A sample from this block of business was taken out near the end of 1946 for another purpose by counting through the renewal cards and picking every 50th card. The sample renewal cards were microfilmed in order to return them to the active file immediately. The exploratory valuation sample of 1,015 cards was obtained from the microfilm file by running it through the viewing machine and stopping the film at every second turn of the crank. The facts from the projection in view were copied by hand on a separate card for this study.

The total sample of 1,015 cards was used to determine an average reserve per $1 of insurance by valuing the individual cards seriatim and adding both the insurance and the reserve on the cards to form two totals. The quotient gave 0.21515 for the average reserve per $1 of insurance. Applying this average to the actual total insurance amounting to $9,587 million, an estimated reserve of $2,063 million was obtained which is $53 million in excess of the true reserve. This gives 0.69 as the ratio of the error to the standard deviation of $77 million estimated from the sample. Hence by this method of sampling, results to this degree of accuracy or better would be obtained in about half of repeated applications.

Individual policies, rather than grouped frequencies from the scatter diagram, were used for the numerical work because of the desire to experiment with various subdivisions. In order to see how the averages would behave in repeated samples, the sample of 1,015 cards was further subdivided into ten groups, each group consisting of all the cards with policy numbers ending in the same digit. Table II, giving the results from this experiment, shows the average reserves per dollar of insurance, the total reserve for this block of business estimated from each of the ten sample averages, and the resulting errors obtained by comparison with the actual reserve.

The actual reserve from the regular arithmetical calculation amounted to $2,010 million. The best estimaté from the small sample of 111 cards was $34 million under the true value, while the worst estimate obtained from a sample of 93 cards was $418 million over the true value. It would be expected that 5 of the ratios of error to the standard deviation would be under 0.6745, disregarding signs, and that 5 would

TABLE II

BEHAVIOR OF AVERAGES FROM REPEATED SAMPLES

Final Digit	Number Policies in Sample	Average Reserve per $1 Insurance	Reserve Estimated from Sample Average and Actual Insurance*	Error (Estimate less Actual)	Ratio of Error to Standard Deviation†
			(Millions)	(Millions)	
0	105	.2279	$2,185	175	0.73
1	93	.2532	2,428	418	1.65
2	109	.1952	1,871	−139	−0.59
3	111	.2062	1,976	− 34	−0.15
4	108	.2517	2,413	403	1.71
5	81	.1959	1,878	−132	−0.49
6	118	.2158	2,068	58	0.26
7	110	.1898	1,820	−190	−0.82
8	87	.2372	2,274	264	1.01
9	93	.1712	1,642	−368	−1.45
True Values		.20966	$2,010		

* Actual Insurance $9,586.83 million.

† Standard Deviation estimated from total sample of 1,015 policies adjusted for size of individual samples.

exceed this value; actually there were 4 under and 6 over. Considering 20 per cent frequency intervals from the lower limit to the upper limit of the distribution, there were 1, 3, 1, 2, 3 ratios in each interval compared with the expected 2.

The subdivision of the sample of 1,015 cards into ten random subgroups illustrates that even with a sample of the order of 100 cards, it is possible to estimate the reserve, and to predict the range of errors with reasonable certainty because the errors, measured in terms of their expected standard deviation, appear to be distributed in reasonably close agreement with the normal frequency curve.

EXPERIMENTAL STRATIFICATION

In the above computation, the policies in the sample were grouped together. It would be natural to inquire if the result might be improved by stratifying the sample.

It is easy to select a valuation sample stratified by years of issue since policy cards usually are filed in numerical order. This breaks them approximately by years of issue. Any other stratification, such as by plans or groups of plans, would be difficult to handle if the cards had to be selected initially on that basis. With some sacrifice of potential efficiency, however, the sample might be stratified after selection for application against such control totals.

An experiment was tried with stratification by year of issue with all plans combined in each year of issue. Average reserve factors per $1 of insurance from the sample were applied to the total business by years of issue. This method of estimating the reserve by a stratified sample reduced the standard deviation from $77 million to $47 million, and reduced the error from $53 million to $19 million, giving a ratio of error to standard deviation of 0.41. This was the best result produced from this sample. The standard deviation of this estimate indicates that the error would be within about $32 million half of the time from repeated samples of 1,015 policies selected in this way. There would be a 5 per cent risk of exceeding an error of $92 million. The results from this experiment are shown in some detail in Table III.

It is interesting to note that, in the last column showing the ratios of the errors to the corresponding standard deviations, there are 12

TABLE III

EXPERIMENTAL SAMPLE STRATIFIED BY YEAR OF ISSUE

Year of Issue	Number of Policies in Sample	Average Reserve per $1 Insurance From Sample	Actual Total Insurance in Force (Millions)	Reserve (Millions)			Estimated* Standard Deviation (Millions)	Ratio Error to Standard Deviation
				Estimated From Sample	Actual	Error		
1945	71	.04146	$775.44	32.15	29.30	2.85	4.07	.70
1944	60	.05718	663.29	37.93	37.11	.82	4.08	.20
1943	37	.06319	540.93	34.18	37.21	− 3.03	6.66	− .46
1942	55	.08048	438.58	35.30	36.56	− 1.26	4.33	− .29
1941	82	.11578	595.02	68.89	61.08	7.81	10.37	.75
1940	59	.12925	440.67	56.96	54.19	2.77	4.88	.57
1939	59	.13596	425.47	57.85	59.33	− 1.48	7.43	− .20
1938	48	.13875	408.94	56.74	67.40	−10.66	11.23	− .95
1937	59	.17848	476.37	85.02	87.15	− 2.13	8.97	− .24
1936	42	.16573	439.46	72.83	90.42	−17.59	5.98	−2.94
1935	45	.22862	418.48	95.67	95.60	.07	7.23	.01
1934	50	.20800	463.69	96.45	112.15	−15.70	13.53	−1.16
1933	58	.26810	374.42	100.38	96.10	4.28	7.70	.56
1932	44	.29552	343.81	101.60	98.40	3.20	11.21	.29
1931	34	.27207	301.41	106.49	118.18	−11.69	9.14	−1.28
1930	25	.38306	353.53	135.42	116.37	19.05	19.55	.97
1929	32	.37959	363.52	137.99	129.65	8.34	11.57	.72
1928	29	.42777	315.03	134.76	121.69	13.07	11.63	1.12
1927	33	.44542	269.53	120.05	112.53	7.52	10.67	.71
1926	16	.33777	185.96	62.81	64.62	− 1.81	3.82	− .47
1925	9	.37549	177.88	66.79	65.47	1.32	4.10	.32
1924	17	.42269	151.01	63.83	57.86	5.97	4.49	1.33
1923	7	.33840	140.76	47.63	56.66	− 9.03	3.89	−2.32
1922	6	.56092	87.57	49.12	36.38	12.74	3.25	3.92*
1921 & prior	38	.49779	346.04	172.25	168.43	3.82	18.80	.20
Total Sample	1,015		$9,586.81	$2,029.09	$2,009.84	19.25	47.17	.41

* Note that the estimate of the standard deviation becomes very unreliable when the number of policies is small; for example there are only 6 policies in the sample for the issue of 1922.

below and 13 above the absolute value 0.6745, compared with the expected 12.5 in each group. If the 25 ratios are classified into 20 per cent frequency intervals from the lower to the upper limit of the distribution, the successive intervals contain 5, 3, 5, 8, 4 ratios as compared with the 5 expected in each interval. This agreement seems satisfactory.

It will be noted that the stratification experiment was based entirely on the sample drawn before considering stratification. Consequently this is not an indication of the best result that might be attained from efficient stratification. If the sample had been drawn originally with a view to stratification, it would have been selected in such a way as to minimize the standard deviation of the final estimate. This would have required drawing more cards in some of the years of issue where the reserve is heavy and a smaller number of cards in other years of issue where the reserve is very light. For example, there were 71 cards in the issue of 1945 where the total reserve is only $29 million but only 32 cards in the issue of 1929, which has a reserve of $130 million. Consequently the issue of 1945 contributed only 17 million million to the total variance, while the issue of 1929 contributed 134 million million to the total variance.

For a given total number of cards in the over-all sample, the selection by years of issue should be made on a basis to minimize the variance of the resulting estimate. The variance was recomputed approximately on the basis of the most efficient distribution of the sample by years of issue, assuming that the estimated variances and correlation coefficients among individual policies would not change within the years of issue. A new value of 1,692 million million was obtained for comparison with the 2,225 million million in the actual sample. These figures indicate that for this distribution a given degree of accuracy could be attained with about one-quarter fewer cards in the sample if it were stratified properly by years of issue to minimize the variance. The corresponding standard deviations would be $41.1 million compared with $47.2 million.

On the basis of stratification by year of issue, but without reflecting the most efficient distribution of the sample, Table IV has been prepared for Metropolitan Regular Ordinary business, giving the approximate number of policies that would be required in a sample in order to provide various degrees of accuracy and risk. The amount of error is shown down the left margin both as a percentage and as an approximate amount of money based on the $2.0 billion reserve in this block of business. The amount of risk is shown across the top of the table. For example, if a sample of 80,000 policies were drawn time after time,

TABLE IV

APPROXIMATE SIZE OF SAMPLE FOR VARIOUS DEGREES OF ACCURACY AND RISK*

Error in Reserve		Risk of Exceeding Designated Error				
Amount (Millions)	%	50%	10%	5%	1%	0.1%
		Approximate Number of Policies Required in Sample*				
$ 2	0.1	230,000	1,100,000	1,400,000	1,950,000	2,500,000
5	0.25	39,000	220,000	310,000	500,000	760,000
10	0.5	9,700	57,000	80,000	140,000	220,000
20	1	2,500	14,500	21,000	35,000	57,000
40	2	610	3,600	5,100	8,800	14,500
80	4	150	900	1,300	2,200	3,600

* Adjusted for sampling from finite universe of 4.27 million Regular Ordinary policies in the Metropolitan Life Insurance Company.

and if the reserve were estimated from each such sample, the error would not exceed 0.5 per cent or $10 million in more than 5 per cent of such estimates. It might be said also that the error would not exceed 0.25 per cent or $5 million in more than one-third of such estimates.

CURRENT APPLICATIONS

The exploratory work reported above gives an indication of the possibilities of sampling procedures in reserve valuations. The question naturally arises as to what extent these techniques are actually in use

As a matter of fact, sampling has been used very little, if at all, for any major reserve valuation but it has been used in a number of instances for smaller items which would be very troublesome to obtain from a complete inventory. One large item of this nature is the deferred premium asset for Metropolitan Regular Ordinary business. This asset arises from the fact that a very considerable proportion of premiums are paid semi-annually, quarterly, or monthly. Since the reserve liability is computed for simplicity on the assumption that all premiums are paid annually, it is necessary to calculate an adjusting item for the balance sheet.

The item of deferred premiums depends entirely on the distribution of premiums according to the month of issue date and according to the mode of payment (annual, semi-annual, quarterly or monthly). Obviously this distribution does not change very rapidly with time, and the results from a single sample can be carried forward for a few years. The problem is to obtain from a sample the estimated average deferred premium at the end of the calendar year per dollar of gross annual premium in force, together with the variance of the estimate. A 2 per cent random sample amounting to 91,011 cards was drawn for the most recent estimate in 1946. For this sample, the average deferred premium per dollar of gross annual premium in force was $0.1964, with the standard deviation indicating an error range of about 1 per cent of the average at the two sigma level. Thus, in an item of $50 million, there is a risk that about once in 20 years, on the average, the error in the estimate will be more than $0.5 million. Although the valuation of this one item involves a very substantial amount, it is very small in relation to the reserves for the total business of the Company.

Another end-of-year adjustment item, "Due and Unpaid Premiums," arises from premiums not paid on the due date because policies are not transferred to the non-forfeiture accounts until sometime after expiration of the "grace" period. This item is much less stable than the deferred item because it depends considerably on current economic psychology in addition to the proportion of gross annual premium due in November and December. The average premium due and unpaid at the end of 1946 amounted to $0.0496 per dollar of gross annual premium in force, giving a total amount of about $12 million. Because this item may change rapidly with time, and is likely to be influenced by Christmas buying, a new estimate is required each year at the end of December. The amount of premiums due and unpaid at the end of 1947 was estimated from a sample of 100 district averages selected from the finite universe of 787 district averages in the total Company. Sampling among these averages rather than among individual policies is much less efficient statistically, but it has been adopted because of certain filing advantages. For the same accuracy, about ten times as many cards are required to obtain the sample district averages as would be used in a random sample of individual policies.

The Metropolitan uses a sampling method for the reserve of nearly $13 million on Additional Insurance purchased by the application of dividends. A sample is taken of all the policies with issue date anniversaries in two different months of the year, giving about 35,000 policies with Additional Insurance. For each of nine subdivisions, the amount of Additional Insurance and the amount of reserve on this insurance is obtained from the sample. Then the average reserve per dollar of insurance is obtained. The total reserve in the nine subdivisions is estimated by multiplying corresponding averages by the total amount of Additional Insurance for each subdivision. The result from one sample is carried forward from one year to the next by an accumulation process which reflects the new dividends applied as single premiums to purchase Additional Insurance, the interest required on the reserve, and adjustments for terminations. The sampling process is repeated every three to five years as a check and correction on the accumulation process.

GENERAL CONSIDERATIONS

A question will arise as to why sampling methods have not been used more widely for valuation purposes, and to what extent such methods may be used in the future.

The answer hinges on the degree of accuracy demanded in the result. With respect to reserve valuation the experimental results presented indicate that a sample of very reasonable size, 21,000 policies for example, would give substantial assurance that the estimated result will be accurate to 1 per cent, but a 1 per cent error in the reserve of $2 billion on the Regular Ordinary business becomes a $20 million error. Neither Company management nor State Insurance Departments would permit errors of this size. Although an error of $20 million is clearly larger than could be tolerated in practice, the amount of error might be reduced within certain limits by increasing the size of the sample. In any sampling plan it is necessary to determine the tolerance limits before proceeding to determine the size of the sample.

The reserve is based on very broad assumptions even if it is computed arithmetically with complete correctness, the assumptions being in regard to the probabilities of death and in regard to the rate of interest. A very moderate change in one of these assumptions may make a very substantial absolute change in the total reserve even though the percentage change may be relatively small. For example, if it is decided to assume that the Company will earn an interest rate of 0.25 per cent less in the future than previously had been assumed, this change in assumption may well add some $60 million to a reserve of $2 billion. Considering the approximate nature of the basic assumptions underlying the reserve calculation, a statistician might well inquire whether the degree of arithmetical accuracy now demanded is really necessary. There are other important practical considerations, however, which must be taken into account. Let us inquire into the source of the requirement and into the extent to which such accuracy is really needed.

Section 205 of the current New York Insurance Law provides in part, "The Superintendent shall annually value, or cause to be valued, the reserve liabilities (hereinafter called reserve) for all outstanding life insurance policies and annuity and pure endowment contracts of every life insurance company doing business in this state, . . . In calculating such reserves, he may use group methods and approximate averages for fractions of a year or otherwise." This last sentence may be interpreted to permit the use of sampling techniques. In practice, however, the New York Department of Insurance is reluctant to approve approximate methods of valuation unless it can be demonstrated that the error is small. Moreover, there is a tendency to consider the error as an absolute amount rather than as a percentage.

After the Company makes the valuation required by law, it forwards its working papers to the State Insurance Department for further checking and review. The formal transmittal to the New York Insurance Department is accompanied by an affidavit signed by the President, the Vice-President and Chief Actuary, and the Assistant Actuaries in immediate charge of the work, and reads in part:

> "That the valuation schedules, summaries and other valuation data . . . have been submitted to the State of New York Insurance Department as the sole basis for calculating said Company's valuation results and issuing its Certificate of Valuation;
>
> "That such schedules have been properly and correctly prepared from the policy and annuity valuation records and that such records accurately and completely cover all of the policies and annuities of the said Company outstanding and paid-for on the 31st day of December, 19 . . . ;
>
> "That such schedules contain, (1) complete and correct groupings according to age, issue, duration, kind and amount of all such policies and annuities, with all the data necessary for making a mean, grouped valuation of the same upon the net premium basis as required by law, and (2) full and complete information necessary for calculating all special net reserve liability items, including reserves for disability and accidental death benefits;
>
> "That the data for the said mean, grouped valuation produces results equal to those required by the terms of the policies and annuities and the law applicable thereto:
>
> "That the above named Actuary has charge of the Company's policy and annuity valuation records and of the valuation schedules and summaries, and has supervised the preparation of the schedules and of the Analysis of Valuation hereto attached and bearing his signature, and that the checks placed on the same insure their accuracy; and
>
> "That the total amount of reserves on said policies and annuities is correct as reported on such schedules, summaries and other valuation data and as set out in the said Analysis of Valuation, according to their best information, knowledge and belief."

It is very difficult to base a complete valuation on sampling methods and still submit such a specific affidavit even if the sample were increased to an impractical size sufficient to eliminate almost all uncertainty.

Aside from the legal requirements, Company management could not tolerate an error of estimation which would become significant when carried into the surplus, and which would become more significant in the surplus earnings of a single year. An error of $1 million in the reserve, in opposite directions at the beginning and at the end of the year, would result in an error of $2 million in the surplus earnings of the single year. The amount of annual surplus earnings, the starting point for the distribution of dividends to policyholders, is about $80 million for the block of business on which the experimental results of this discussion were based. It is clear that the amount of acceptable error due to sampling fluctutions is limited severely by the necessity for accuracy, consistency, and continuity in this important item.

FUTURE APPLICATIONS

It appears likely that the immediate extension of sampling techniques in reserve valuations will be limited to the reserves on related benefits, including certain paid-up policies, on which the total reserve is so small that an error of the order of 1 per cent does not look too large in relation to the surplus earnings in a single year.

One can only speculate in regard to the future development of sampling techniques for the more important valuation of the main life insurance reserves, running into billions. Application to independent checking and auditing might well be a possible development of sampling techniques. State Insurance Departments frequently consider the problem of reconciling the active file of renewal cards with the separate book record in force classification accumulated over the years from the transactions and used as the basis for the arithmetical calculation of the reserve liability. The cost of a complete inventory would be excessive, not to mention the difficulty of "freezing" the very active file of renewal cards long enough to carry out the work. The exploratory methods used in this discussion, with some extension and with a reasonably large sample, could provide an independent check on the total reserve which would detect any error of sufficient size to be significant when carried into surplus.

Progress may be expected in developing sampling techniques for the main life insurance reserve, perhaps by the indirect approach of sampling transactions rather than by direct frontal attack. It probably will take years of study, however, before sampling methods will become practicable and generally acceptable for the main life insurance reserve liability in the Annual Statement.

RECENT DEVELOPMENTS IN GRADUATION AND INTERPOLATION*

THOMAS N. E. GREVILLE
National Office of Vital Statistics, U. S. Public Health Service

Some developments during the 10-year period 1938–1947 in the techniques of graduation, or the smoothing of data, and of interpolation, are reviewed and summarized. In preparing this summary of developments in graduation and interpolation during the past decade, it has been the author's intention to refer to the more important results which have come to his attention, and which he considers are likely to be of interest to practical statisticians and actuaries who have occasion to perform graduations or interpolations of actual data. No attempt has been made to deal with the more abstract topic of interpolatory function theory. Within the limits thus set, there has been no intention to slight anyone, and the author will appreciate having his attention called to any omissions. For derivations, proofs, and more detailed treatment, the references given in the bibliography may be consulted.

I. GRADUATION

Characteristic function of a graduation formula. Graduation (or "smoothing") has been defined as "the process of securing from an irregular series of observed values of a . . . variable a smooth regular series of values consistent in a general way with the observed series of values" which "is then taken as a representation of the underlying law which gave rise to the series of observed values" [17].[1] A traditional method of graduation is based on the use of weighted moving averages.[2] This method may be represented by a formula of the form:

$$F_n = \sum_{\nu=-\infty}^{\infty} y_\nu L_{n-\nu} \tag{1}$$

where y_ν denotes the sequence of given data to be smoothed, F_n is the "smoothed" value of y_n and the numbers L_ν are the numerical coefficients which characterize the formula. In practice, of course, both the given data and the coefficients L_ν are limited to a finite range.

In an elegant study of this method of graduation, Schoenberg [18] has defined a "characteristic function" for each such smoothing formula, and has shown that the properties of the formula are uniquely related to certain properties of this function, which thus uniquely "characterizes" the smoothing formula; so that, knowing the function, one can deduce the formula, and vice versa. Specifically the characteristic function of the given sequence is defined as

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

[1] Numbers in brackets refer to the bibliography at the end of this article.

[2] This method is also referred to as graduation by "adjusted averages" or by "linear compounding." Prominently associated with its early development are the names of De Forest [28] and Sheppard [19, 20].

$$T(u) = \sum_{n=-\infty}^{\infty} y_n e^{inu},$$

where u is an arbitrary parameter. Similarly the characteristic function of the smoothing formula is defined as

$$\phi(u) = \sum_{n=-\infty}^{\infty} L_n e^{inu}.$$

By multiplication, it is evident that

$$T(u)\phi(u) = \sum_{n=-\infty}^{\infty} F_n e^{inu},$$

so that the characteristic function of the "smoothed" sequence is the product of the characteristic functions of the given data and of the smoothing formula. If the coefficients L_ν are symmetrical in the sense that $L_{-\nu} = L_\nu$, the characteristic function of the formula can be expressed in the form of the Fourier series:

$$\phi(u) = L_0 + 2 \sum^{-} L_n \cos nu.$$

In terms of the characteristic function, Schoenberg proposes a criterion for stating when a formula of this type is really a "smoothing" formula in the sense that its application may be expected actually to increase the smoothness of the data. It has been customary to judge the smoothness of a graduation by the size of some specified order of finite differences of the graduated values. Using the same approach, he establishes the relations

$$\sum_{n=-\infty}^{\infty} (\Delta^m y_n)^2 = \frac{1}{2\pi} \int_0^{2\pi} (2 \sin \tfrac{1}{2}u)^{2m} \,|\, T(u) \,|^2 du,$$

$$\sum_{n=-\infty}^{\infty} (\Delta^m F_n)^2 = \frac{1}{2\pi} \int_0^{2\pi} (2 \sin \tfrac{1}{2}u)^{2m} \,|\, T(u) \,|^2 \,|\, \phi(u) \,|^2 du,$$

where m is any nonnegative integer. Observing that the integrands of these two expressions differ only by the factor $|\phi(u)|^2$, he defines a smoothing formula as one for which $|\phi(u)| \leqq 1$, for $0 \leqq u \leqq 2\pi$, with the additional stipulations[3]

$$\sum_{n=-\infty}^{\infty} L_n = 1 \quad \text{and} \quad \sum_{n=-\infty}^{\infty} |L_n| < \infty.$$

Application of orthogonal polynomials to graduation formulas of maximum smoothness. A smoothing formula is said to be exact for the degree m if it reproduces exactly the values of any polynomial of degree not exceeding m. In this connection, Schoenberg shows that a formula of the type (1) is exact for the degree m if and only if $\phi(u)-1$ has at $u=0$ a zero of order $m+1$, or, in other words, if this expression and its first m derivatives all vanish for $u=0$.

If $L_{-n}=L_n=0$ for $n>k$, where L_k and L_{-k} are not zero, the formula is said to be of span $2k+1$. If each of $2k+1$ consecutive ungraduated values is assumed subject to a random error with zero mean and variance the same for all the $2k+1$ values, then the ratio of the mean squared error of the graduated value to that of the ungraduated value is

$$\sum_{n=-k}^{k} L_n^2.$$

Therefore particular interest attaches to those formulas which minimize this ratio for a given span and degree reproduced. These are said to effect the maximum "reduction of error," and are sometimes referred to as formulas of "maximum weight."[4] Those formulas for which

$$\sum_{n=-k-3}^{k} (\Delta^3 L_n)^2$$

is a minimum have frequently been called formulas of maximum smoothness. However, it must be admitted that the choice of third differences for this purpose is somewhat arbitrary, and that an argument for the use of some other order of differences could be advanced with equal justification. It has long been known [7, 1, 16] that the computation involved in performing a graduation by a formula of maximum error reduction can be considerably simplified by the use of a set of orthogonal polynomials associated with the name of Chebyshev.

---

[3] Schoenberg limits the definition to formulas with symmetric coefficients, but it appears to be equally applicable to those with unsymmetric coefficients.

[4] This is, of course, a different use of the word "weight" from that employed when a "weighted average" is referred to. For a fuller discussion of the principles underlying these formulas and the formulas of maximum smoothness next mentioned, see [17, Chapter 4].

The Chebyshev polynomials $T_i(x, n)$ are given by the formula:

$$T_i(x, n) = \Delta^i[x_{(i)}(x-n)_{(i)}],$$

where $x_{(i)}$ denotes the expression $x(x-1)\cdots(x-i+1)/i!$ These polynomials are orthogonal over the summation range 0 to $n-1$: in other words, we have:

$$\sum_{x=0}^{n-1} T_i(x, n)T_j(x, n) = 0$$

if $i \neq j$, while

$$\sum_{x=0}^{n-1} [T_i(x, n)]^2 = S_i(n) \qquad \frac{n(n^2-1)(n^2-4)\cdots(n^2-i^2)}{(2i+1)(i!)^2}$$

If we now consider the smoothing formula for obtaining a graduated value of y_{a+x} which involves the given values y_a to y_{a+n-1}, inclusive, is exact for the degree r, and gives maximum reduction of error, it is well known that the coefficient of y_{a+z} is given by

$$L_{z-x} = \sum_{i=0}^{r} \frac{T_i(x, n)T_i(z, n)}{S_i(n)}. \tag{2}$$

This remarkably simple expression for the coefficient results from the orthogonal property of the polynomials $T_i(x, n)$. As fairly extensive tables of the values of $T_i(x, n)$ and $S_i(n)$ have been published [3], this formula makes it easy to compute linear compound coefficients, not only for the symmetrical "mid-panel" formulas used over the greater part of the range of given data, but also for unsymmetrical "end-panel" formulas to be used near the ends of the range. Greville [13] has recently shown that the coefficients for the formulas of maximum smoothness also can be expressed in terms of the Chebyshev polynomials. Using the same notation as in formula (2), with the addition that m denotes the order of differences taken as a criterion of smoothness, the coefficient of y_{a+z} is[5]

$$L_{z-x} = (-1)^m(z+m)^{(m)}(z-n)^{(m)} \sum_{i=m}^{r+m} \frac{\Delta^m T_i(x, n+m)\Delta^m T_i(z, n+m)}{(m+i)^{(2m)}S_i(n+m)}$$

where $z^{(m)}$ denotes $z(z-1)\cdots(z-m+1)$.

Difference equation graduation. For a number of years, a favorite method of graduation among American actuaries has been the "difference equation" method developed by Whittaker [27] and Henderson [14, 15]. This method results from minimizing the expression:

---

[5] This formula represents a slight improvement over the one given in the aforementioned article. For derivation of this formula, see the author's reply to the discussion of [13], *Record*, American Institute of Actuaries, Vol. 37, Part I, No. 75, April 1948.

$$\sum_{x=a}^{b} (u_x - y_x)^2 + k \sum_{x=a}^{b-m} (\Delta^m u_x)^2,$$

where y_x denotes the given values, which are assumed to be known over the range $x=a$ to $x=b$, u_x denotes the graduated values, m is a specified order of differences, and k is a parameter to be chosen by the graduator. The first term in the above expression is a measure of fidelity to the given data, while the second is a measure of smoothness, and the parameter k enables the graduator to assign whatever relative weight he desires to these two requirements. It turns out that the graduated values satisfy the difference equation:

$$u_x + (-1)^m k \delta^{2m} u_x = y_x, \tag{3}$$

where δ denotes a difference taken centrally. This equation holds for $a \leqq x \leqq b$, with the understanding that $\Delta^m u_x = 0$ for $x < a$ and for $x > b - m$. Early methods of solving the difference equation (3) subject to these end conditions produce only approximations to the graduated values near the beginning of the sequence, or involve successive approximation [14, 15], or else involve extensive computation [21]. Weaver [25] published in 1943 an abbreviated method of solution which gives the exact values directly. His method is based, in essence, on modern "short-cut" methods of solving simultaneous linear equations [8]. Let

$$v_x = cy_x + c_{1,x-1}v_{x-1} + c_{2,x-2}v_{x-2} + \cdots + c_{m,x-m}v_{x-m} \tag{4}$$

$$u_x = c_{0,x}v_x + c_{1,x}u_{x+1} - c_{2,x}u_{x+2} + \cdots + (-1)^m c_{m,x}u_{x+m}, \tag{5}$$

where v_x denotes a sequence of "partly graduated" values intermediate between u_x and y_x, the coefficient c of y_x is a constant, and it is understood that those "c" coefficients are zero which would be applied to "v's" with subscripts less than a or to "u's" with subscripts greater than b. Expanding the differences in equation (3) in terms of the u_x values themselves gives a set of simultaneous equations connecting the "u's" and "y's." Another such set of equations is obtained by solving equation (5) for v_x and substituting the resulting expression for all the "v's" in equation (4). Equating coefficients in the two sets of equations gives a set of relationships between the "c" coefficients from which these coefficients can be successively determined. The possibility of expressing the relationship between the "u's" and "y's" in the form of equations (4) and (5) depends on the fact that the matrix of the coefficients in the simultaneous equations is symmetrical about the principal diagonal. Spoerl [23] has previously described the application of this method to the solution of simultaneous equations derived from difference equations.

Spoerl [22] has also generalized the difference equation method of graduation by considering the so-called "mixed difference case," in which the expression to be minimized is taken as

$$\sum_{x=a}^{b} (u_x - y_x)^2 + h \sum_{x=a}^{b-2} (\Delta^2 u_x)^2 + k \sum_{x=a}^{b-3} (\Delta^3 u_x)^2,$$

where h and k are both arbitrary parameters.

II. INTERPOLATION

Systematization of osculatory interpolation formulas. One of the most common practical problems in interpolation is that which is sometimes called the problem of subtabulation. This is illustrated by the case in which values of the rate of mortality, or some other actuarial function depending on age, are given for every fifth or tenth year of age, and it is desired to supply the intermediate values for each single year of age. The traditional method of interpolation by means of a moving polynomial arc of degree $k-1$ determined by k consecutive given values gives rise to troublesome discontinuities when applied to empirical data. This long ago led actuaries to alter the conditions imposed on the polynomial arcs by decreasing the number of given points which the curve is required to pass through, but requiring the continuity of first derivatives, and, in some cases, higher derivatives, of the composite interpolation curve [29]. Greville [10] gave, in 1944, a systematization of the formulas of this type, which enables one to develop rapidly the formula, if any, satisfying given predetermined conditions as to span, orders of derivatives required to be continuous, degree of the polynomial arcs used, exactness to a specified order of differences, and reproduction of the given values.

For example, it is shown that a linear compound formula for interpolation in equal intervals by means of a composite interpolation curve made up of spliced polynomial arcs and having a continuous first derivative can be expressed in the form

$$\begin{aligned} y_{n+x} = {} & [z^2(1+2x)(A_0 - B_1) + zB_1 - \tfrac{1}{2}xz^2\delta^2 B_1]y_n \\ & + [x^2(1+2z)(A_0 - B_1) + xB_1 - \tfrac{1}{2}x^2z\delta^2 B_1]y_{n+1} \\ & + x^2z^2P(y, x) \qquad (6) \end{aligned}$$

where $0 \leqq x \leqq 1$, z denotes $1-x$, δ denotes a difference taken centrally, A_0 and B_1 are finite difference operators of the form

$$c_0 + c_1\delta^2 + c_2\delta^4 + \cdots ,$$

and $P(y, x)$ is any linear compound of the "y's" having polynomials in x as coefficients. Suppose it is desired to obtain a 4-term formula using third-degree polynomial arcs, correct to second differences and reproducing the given values. Since the polynomial arcs are to be of the third degree, we must have $P(y, x)=0$. Reproduction of the given values means that $A_0=1$; to have a 4-term formula, B_1 must be limited to the first term c_0 only; and for correctness to second differences c_0 must have the value unity. Substituting these values in equation (6) gives the well known Karup-King interpolation formula.

On the other hand, if it should be desired to forego the requirement that the given values be reproduced but to require instead that the polynomial arcs be of the second degree, A_0 is of the form $1+k\delta^2$. In order for the polynomials to be of the second degree we must have $A_0=(1+\frac{1}{4}\delta^2)B_1$. Since $B_1=1$, that means that $A_0=1+\frac{1}{4}\delta^2$ which gives

$$y_{n+x} = (z + \tfrac{1}{4}z^2\delta^2)y_n + (x + \tfrac{1}{4}x^2\delta^2)y_{n+1},$$

which is Jenkins' second difference smoothing interpolation formula, a well-known formula which not only interpolates but, at the same time smooths the given data.

Characteristic function of an interpolation formula. Generalizing his previously mentioned study of graduation, Schoenberg [18] has also developed a characteristic function for interpolation formulas expressible in linear compound form, so that the interpolatory properties of the formula can be inferred from certain properties of the characteristic function. Let the interpolation formula be

$$F(x) = \sum_{\nu=-\infty}^{\infty} y_\nu L(x - \nu),$$

where the numbers y_ν represent the given data, $F(x)$ denotes the interpolated value, and $L(x)$ is called the "basic function" of the interpolation formula. The characteristic function[6] is defined as

$$g(u) = \int_{-\infty}^{\infty} L(x)e^{iux}dx.$$

---

[6] A function $g(u)$ defined in this manner in terms of another function $L(x)$ is sometimes referred to as the Fourier transform of $L(x)$.

If, as is usually the case, $L(x)$ is an even function, this may be written in the form

$$g(u) = \int_{-\infty}^{\infty} L(x) \cos uxdx.$$

Schoenberg then shows that the interpolation formula is correct to kth differences if $g(u) - 1$ has a zero of order $k+1$ at $u=0$ and, at the same time, $g(u)$ has zeros of order $k+1$ for $u = 2\pi n$, where n is any integer other than zero. He also shows that the formula reproduces the given values if and only if

$$\sum_{\nu=-\infty}^{\infty} g(u + 2\pi\nu) \equiv 1.$$

Incidentally, he further shows that the basic functions of the well-known interpolation formulas can be expressed in a remarkably compact form through the use of Fourier integrals. For example, Jenkins' well-known fifth-difference smoothing interpolation formula is given by:

$$L(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{\sin \frac{1}{2}u}{\frac{1}{2}u} \right)^2 \left(\tfrac{4}{3} - \tfrac{1}{3} \cos u\right) e^{iux} dx.$$

Minimized difference interpolation. Beers [4] has suggested that, as a practical matter, in most actuarial and statistical applications the interpolator usually is interested in interpolated values only for certain discrete, equally spaced arguments. For example, in the case of actuarial functions depending on age, it is usually only the values for integral ages that are of importance. Accordingly, he argues that the consideration of continuous curves with continuous derivatives is irrelevant and unnecessary, and proposes the development of tables of interpolation coefficients in linear compound form designed to minimize, on the average, the sum of the squares of a specified order of finite differences of the interpolated values, subject to stated requirements as to span, correctness to a specified order of differences, and reproduction of given values. Thus, in his method the interpolated points are not thought of as lying on any particular polynomial or other curve, and no mention is made of derivatives or their continuity.

For example, in the case of a 6-term formula correct to fourth differences and reproducing the given values, if it is assumed that any two consecutive fifth differences of the *given* data are independently distributed random variables with zero mean and equal variances, the interpolation coefficients for subtabulation in fifths which minimize, on

TABLE I

INTERPOLATION COEFFICIENTS FOR BEERS' MINIMIZED FIFTH-DIFFERENCE FORMULA FOR SUBDIVIDING THE INTERPOLATION INTERVAL INTO FIFTHS

PART A—TO BE USED FOR THE FIRST TWO INTERVALS

x	Coefficients of y_x To Obtain:							
	$y_{.2}$	$y_{.4}$	$y_{.6}$	$y_{.8}$	$y_{1.2}$	$y_{1.4}$	$y_{1.6}$	$y_{1.8}$
0	+.6667	+.4072	+.2148	+.0819	−.0404	−.0497	−.0389	−.0191
1	+.4969	+.8344	+1.0204	+1.0689	+.8404	+.6229	+.3849	+.1659
2	−.1426	−.2336	−.2456	−.1666	+.2344	+.5014	+.7534	+.9354
3	−.1006	−.0976	−.0536	−.0126	−.0216	−.0646	−.1006	−.0906
4	+.1079	+.1224	+.0884	+.0399	−.0196	−.0181	−.0041	+.0069
5	−.0283	−.0328	−.0244	−.0115	+.0068	+.0081	+.0053	+.0015

PART B—TO BE USED EXCEPT FOR THE FIRST TWO AND LAST TWO INTERVALS

	Coefficients of y_x To Obtain:			
	$y_{n+.2}$	$y_{n+.4}$	$y_{n+.6}$	$y_{n+.8}$
$n-2$	+.0117	+.0137	−.0087	+.0027
$n-1$	−.0921	−.1101	+.0771	−.0311
n	+.9234	+.7194	+.4454	+.1854
$n+1$	+.1854	+.4454	+.7194	+.9234
$n+2$	−.0311	−.0771	−.1101	−.0921
$n+3$	+.0027	+.0087	+.0137	+.0117

PART C—TO BE USED FOR THE LAST TWO INTERVALS

x	Coefficients of y_x To Obtain:							
	$y_{z-1.8}$	$y_{z-1.6}$	$y_{z-1.4}$	$y_{z-1.2}$	$y_{z-.8}$	$y_{z-.6}$	$y_{z-.4}$	$y_{z-.2}$
$z-5$	+.0015	+.0053	+.0081	+.0068	−.0115	−.0244	−.0328	−.0283
$z-4$	+.0069	−.0041	−.0181	−.0196	+.0399	+.0884	+.1224	+.1079
$z-3$	−.0906	−.1006	−.0646	−.0216	−.0126	−.0536	−.0976	−.1006
$z-2$	+.9354	+.7534	+.5014	+.2344	−.1666	−.2456	−.2336	−.1426
$z-1$	+.1659	+.3849	+.6229	+.8404	+1.0689	+1.0204	+.8344	+.4969
z	−.0191	−.0389	−.0497	−.0404	+.0819	+.2148	+.4072	+.6667

the average, the sum of the squares of the fifth differences of the interpolated values are those given in Table I. While, on theoretical grounds, a fairly high negative correlation between successive fifth differences would be expected, it will commonly be found in practice that the actual behavior of the fifth differences of empirical data approximates closely enough to the assumptions stated so that Beers' formula yields plausible results with a highly satisfactory degree of smoothness.

Tables of interpolation coefficients deduced on the same or similar principles have been published by Beers [4, 5, 6], Greville [11], and Wells [26].

Relationship of interpolation to summation graduation. Vaughan [24] has shown that there is an intimate connection between linear compound subtabulation and the "summation" formulas of graduation which have received much attention from English actuaries. The latter are a special type of linear compound graduation formula, in which the relationship between the graduated and ungraduated value is of the form:

$$u_x = \frac{[p][q][r]\cdots}{pqr\cdots} f(\delta) y_x,$$

where $p, q, r \cdots$ are positive odd integers, $[p]$ is an operator denoting the summation of p consecutive given values symmetrically situated about the value to which the operation is applied, and $f(\delta)$ is an "operand" of the form $1+c_1\delta^2+c_2\delta^4+\cdots$, which is usually chosen so as to make the graduation formula correct to a specified order of differences. This type of formula was of greater interest and importance before computing machines came into general use, as the use of summations in the formulas then resulted in a substantial short cut in computation.

Let a subtabulation formula for dividing the intervals between given data into k parts be

$$u_x = \sum_{\nu=-\infty}^{\infty} L_{x-\nu k} y_{\nu k}.$$

Then, there corresponds uniquely to this interpolation formula the graduation formula

$$U_x = \frac{1}{k} \sum_{\nu=-\infty}^{\infty} L_{x-\nu} Y_\nu.$$

Vaughan shows that the latter formula can be expressed in the form

$$U_x = [k]^{n+1} f(\delta) Y_x,$$

where n is the order of differences to which the interpolation formula is correct. Moreover, the terms of $f(\delta)$ as far as δ^n must be such that the entire graduation formula is correct to nth differences. As the number of possible terms in $f(\delta)$ is limited by the span of the interpolation formula, this leaves, in general, only a small number of undetermined coefficients. This principle therefore provides an alternative method of derivation of interpolation coefficients satisfying such criteria as Beers has proposed.

Interpolation by applying a least square criterion over the entire range. Vaughan [24] and Greville [12] have considered independently a somewhat different approach to the subtabulation problem. In this approach, instead of expressing each interpolated value in terms of a few neighboring values, the object is "to ascertain definitely the interpolated terms which will have over the whole series the smallest possible sum of squared differences of any given order" [24]. It is shown that in this case the given and interpolated values are connected by the relation:

$$y_{\nu k+x} = \frac{E^{-x}F(E)}{f(E^k)}\, y_{\nu k}, \tag{7}$$

where k is the interval of subtabulation, E denotes the displacement operator defined by $Ey_x = y_{x+1}$,

$$F(E) = \frac{(E^{k/2} - E^{-k/2})^{2m}}{(E^{1/2} - E^{-1/2})^{2m}} \tag{8}$$

(m being the order of differences to be minimized), and $f(E^k)$ is an expression consisting of only those terms in the expanded quotient of formula (8) in which the exponent of E is a multiple of k. The solution of the difference equation (7) is facilitated by introducing an intermediate sequence v_x which vanishes except when x is an integral multiple of k, such that

$$y_{\nu k+x} = E^{-x}F(E)v_{\nu k},$$

and

$$f(E^k)v_{\nu k} = y_{\nu k}. \tag{9}$$

The difference equation (9) may be solved by a method analogous to that of Weaver previously described, or by expanding the function $1/f(t)$ in a power series of the form: $a_0+a_1(t+t^{-1})+a_2(t^2+t^{-2})+\cdots$.

Interpolation by "quadratic cross-means." Another type of practical interpolation problem is that in which only a single isolated value, or only a few such values, are desired. Probably the best method of estimating such values for empirical data is the classical one of fitting a polynomial of degree k to $k+1$ given values. One of the most effective methods of carrying out the computation for such interpolation is the iterative process given in 1932 by Aitken [2], which has recently been improved by Feller [9], who demonstrates Aitken's basic process in the following manner. Let it be required to compute the value $f(\xi)$ of a polynomial of the nth degree, $f(x)$, given $f_k = f(x_k)$ for $k = 0, 1, \cdots, n$.

We define $$f^{(1)}(x) \quad \begin{matrix} f_0 & x_0 - \xi \\ f(x) & x - \xi \end{matrix} \quad \div (x - x_0)$$

and note that it is a polynomial of degree $n-1$, and that $f^{(1)}(\xi) = f(\xi)$. Hence we are required to find $f^{(1)}(\xi)$ knowing

$$f_k^{(1)} = \begin{matrix} f_0 & x_0 - \xi \\ f_k & x_k - \xi \end{matrix} \quad \div (x_k - x_0)$$

for $k = 1, 2, \cdots, n$. Thus the problem has been reduced to a similar problem in which the number of given values is one less, and in like manner it is further reduced to $n-2$, and so on.

The improvement introduced by Feller arises in interpolating by "quadratic cross-means." The latter process, which originally appeared to work only for an even number of data, can now be equally well adapted to an odd number. However, it is still necessary that the given data be symmetrically placed with respect to some point $x = a$. The essential feature of the method consists in regarding the interpolation polynomial as a polynomial of lower degree in the *square* of the independent variable. This device reduces by about one-half the number of steps in the iterative process. Denote two symmetrically placed points by x_k and x_{-k}. The point $x_0 = a$ is included among the data only if n is even. Consider

$$\phi(x) = \begin{matrix} f(2a - x) & 2a - x - \xi \\ f(x) & x - \xi \end{matrix} \quad \div 2(x - a)$$

$$\psi(x) \quad \begin{matrix} -f(2a - x) & 2a - x - \xi \\ f(x) & x - \xi \end{matrix} \quad \div 2(x - a).$$

Obviously $\phi(x)$ and $\psi(x)$ are even functions of $x - a$, and hence polynomials in $t = (x - a)^2$. Moreover, $\phi(\xi) = \psi(\xi) = f(\xi)$.

If $n = 2m - 1$, the problem is reduced to finding the value for $t = (\xi - a)^2$ of the polynomial of $(m-1)$ th degree $\phi(a + \sqrt{t})$, given its values

$$\phi_k = \begin{matrix} f_{-k} & x_{-k} - \xi \\ f_k & x_k - \xi \end{matrix} \quad (x_k - x_{-k})$$

for $t = (x_k - a)^2$, $k = 1, 2, \cdots, m$.

On the other hand if $n = 2m$, we compute the values

$$\psi_k = \begin{matrix} -f_{-k} & x_{-k} - \xi \\ f_k & x_k - \xi \end{matrix} \quad \div 2(\xi - a)$$

for $t = (x_k - a)^2$, $k = 0, 1, \cdots, m$, and proceed as before.

BIBLIOGRAPHY

[1] Aitken, A. C., "On the Graduation of Data by the Orthogonal Polynomials of Least Squares," *Proceedings of the Royal Society of Edinburgh,* Session 1932–33, Part I, Vol. 53, pp. 54–78.

[2] Aitken, A. C., "On Interpolation by Iteration of Proportional Parts, without the Use of Differences," *Proceedings of the Edinburgh Mathematical Society,* Series 2, Vol. 3, 1932, pp. 56–76.

[3] Anderson, R. L. and Houseman, E. E., *Tables of Orthogonal Polynomial Values Extended to* $N=104$, Research Bulletin 297, pp. 595–672, Agricultural Experiment Station, Iowa State College of Agriculture and Mechanic Arts, Statistical Section, Ames, Iowa, April 1942.

[4] Beers, H. S., "Six-Term Formulas for Routine Actuarial Interpolation," *Record,* American Institute of Actuaries, Vol. 33, Part II, No. 68, November 1944, pp. 245–260.

[5] Beers, H. S., "Modified-Interpolation Formulas that Minimize Fourth Differences," same *Record,* Vol. 34, Part I, No. 69, June 1945, pp. 14–20.

[6] Beers, H. S., "Premium Interpolation," *Transactions,* Actuarial Society of America, Vol. 48, Part I, No. 117, May 1947, pp. 53–75.

[7] Chebyshev, P. L., *Oeuvres,* Vol. 1, pp. 203–230, 381–384, 473–498, 541–560; Vol. 2, pp. 219–242, Académie Impériale des Sciences, St. Petersburg, 1879 and 1907.

[8] Dwyer, P. S., "The Solution of Simultaneous Equations," *Psychometrika,* Vol. 6, No. 2, April 1941, pp. 101–129.

[9] Feller, W., "On A. C. Aitken's Method of Interpolation," *Quarterly of Applied Mathematics,* Vol. 1, No. 1, April 1943, pp. 86–87.

[10] Greville, T. N. E., "The General Theory of Osculatory Interpolation," *Transactions,* Actuarial Society of America, Vol. 45, Part II, No. 112, October 1944, pp. 202–265.

[11] Greville, T. N. E., "Actuarial Note: Some Extensions of Mr. Beers' Method of Interpolation," *Record,* American Institute of Actuaries, Vol. 34, Part I, No. 69, June 1945, pp. 21–34.

[12] Greville, T. N. E., "Subtabulação por Mínimos Quadrados de Diferenças Finitas," *Boletim do Instituto Brasileiro de Atuária,* Ano 2, No. 2, September 1946, pp. 7–34.

[13] Greville, T. N. E., "Adjusted Average Graduation Formulas of Maximum Smoothness," *Record,* American Institute of Actuaries, Vol. 36, Part II, No. 74, October 1947, pp. 249–264.

[14] Henderson, R., "A New Method of Graduation," *Transactions,* Actuarial Society of America, Vol. 25, Part I, No. 71, May 1924, pp. 29–40.

[15] Henderson, R., *Mathematical Theory of Graduation* (Actuarial Studies No. 4), Chapter 4, Actuarial Society of America, New York, 1938.

[16] Lidstone, G. J., "Notes on Orthogonal Polynomials and Their Application to Least Square Methods of (1) Fitting Polynomial Curves to Data, (2) Graduation by Weighted Means," *Journal of the Institute of Actuaries,* Vol. 64, Part II, No. 309 (1933), pp. 128–159.

[17] Miller, M. D., *Elements of Graduation,* pp. 3–4, Actuarial Society of America, New York, 1946.

[18] Schoenberg, I. J., "Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions," *Quarterly of Applied Mathe-*

matics, Vol. 4, No. 1, April 1946, pp. 45–99, and No. 2, July 1946, pp. 112–141.

[19] Sheppard, W. F., "Reduction of Errors by Means of Negligible Differences," *Proceedings of the Fifth International Congress of Mathematicians*, Vol. 2, pp. 348–384, Cambridge, 1913.

[20] Sheppard, W. F., "Reduction of Error by Linear Compounding," *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 221, March 1921, pp. 199–237.

[21] Spoerl, C. A., "The Whittaker-Henderson Graduation Formula A," *Transactions*, Actuarial Society of America, Vol. 38, Part II, No. 98, October 1937, pp. 403–462.

[22] Spoerl, C. A., "The Whittaker-Henderson Graduation Formula A, the Mixed Difference Case," same *Transactions*, Vol. 42, Part II, No. 106, September 1941, pp. 292–313.

[23] Spoerl, C. A., "A Fundamental Proposition in the Solution of Simultaneous Linear Equations," same *Transactions*, Vol. 44, Part II, No. 110, October 1943, pp. 276–288.

[24] Vaughan, H., "Some Notes on Interpolation," *Journal of the Institute of Actuaries*, Vol. 72, Part III, No. 335, June 1946, pp. 482–497.

[25] Weaver, C. L., "Actuarial Note: Modification of Working Formulas in Whittaker-Henderson A Graduation Method to Produce the Moments Automatically," *Transactions*, Actuarial Society of America, Vol. 44, Part I, No. 109, May 1943, pp. 27–30.

[26] Wells, E. H., Discussion of [4], *Record*, American Institute of Actuaries. Vol. 44, Part I, No. 109, May 1943, pp. 27–30.

[27] Whittaker, E. T. and Robinson, G., *The Calculus of Observations*, Second Edition, pp. 303–316, Blackie and Son, London, 1937.

[28] Wolfenden, H. H., "On the Development of Formulae for Graduation by Linear Compounding, with Special Reference to the Work of Erastus L. DeForest," *Transactions*, Actuarial Society of America, Vol. 26, Part I, No. 73, May, 1925, pp. 81–121.

[29] Wolfenden, H. H., *Population Statistics and their Compilation* (Actuarial Studies No. 3), pp. 102–109, Actuarial Society of America, New York, 1925.

SAMPLING ERRORS IN MORTALITY AND OTHER STATISTICS IN LIFE INSURANCE*

DONALD D. CODY

The Equitable Life Assurance Society of the United States

The types, sources, and extent of data entering into the analytical statistics in life insurance companies are outlined. The statistical errors in mortality investigations arising as a result of the limited extent of the death claim experience or arising from the use of exposure sampling techniques are discussed. Comments are also made about the experience rating problem in group insurance and about a possible application of sequential analysis to the checking of files.

TYPES, SOURCES, AND EXTENT OF DATA

ACTUARIES are concerned with a number of different types of statistical rates and probabilities. The most important of these are the mortality rate, the rate of occurrence of permanent disability, the rate of termination of such disability by recovery or death, the accidental death rate, the rate of claim under accident and health policies, the rate of claim under hospitalization and medical reimbursement policies, and the rate of termination of policies. The sources of data for making estimates of these rates and probabilities are the company's files, which, in some cases, are designed primarily for furnishing such data, and in other cases, are designed primarily for other purposes. A description of certain files from which data for life insurance mortality investigations are taken in the Equitable Life Assurance Society will serve as an illustration.

The basic statistical data regarding death for a life insurance mortality investigation come almost exclusively from the death claim file, which is a file of I.B.M. punched cards prepared from the claim records filed when a policy becomes payable by death. A specimen card is shown in Fig. 1. Each calendar year some 16,000 to 18,000 of these cards are added to these files.

The primary source of information concerning the exposure to the risk of death in large over-all mortality investigations of standard medically examined business comes from another file of I.B.M. punched cards which is used primarily for the annual valuation of the company's reserve liability. This file is summarized mechanically to give the numbers and face amounts of insurance on policies in force on December 31

---

* A paper presented at the 107th Annual Meeting of the American Statistical Association, New York City, December 30, 1947.

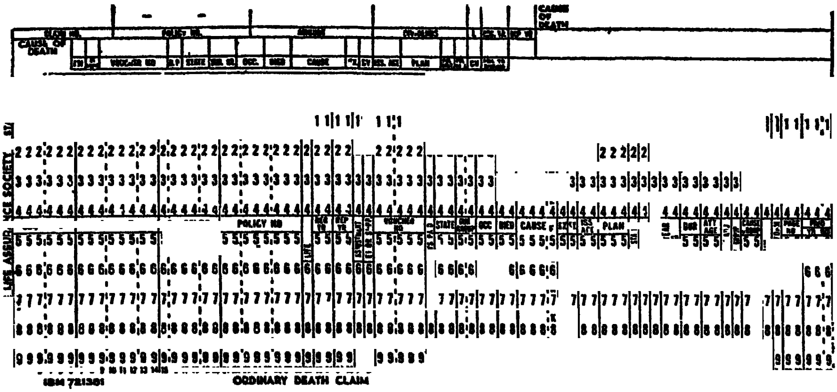


FIGURE 1

DEATH CLAIM DETAIL CARD

of each calendar year; these are classified according to rating group, plan of insurance, calendar year of issue, and age at issue. The data from this file must be adjusted to differentiate policies with some special attribute which makes them non-homogeneous with the bulk of the policies, e.g., policies issued without a medical examination. There are now about 2,000,000 cards in the valuation detail card file of the Equitable, and the file is increased by a net addition of 50,000 to 100,000 cards each year. However, the summary cards (see Fig. 2 for a specimen) actually used number less than 100,000.

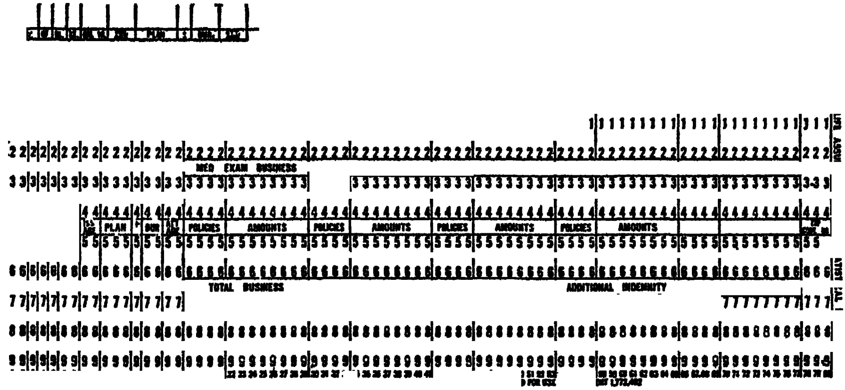


FIGURE 2

SUMMARY CARD—INSURANCE IN-FORCE

It is also necessary to keep a special mortality class file of I.B.M. punched cards, which are made up from application papers at issue, in the case of special policies such as those involving medical impairments and occupational hazards (see Fig. 3). There are about 1,250,000 cards in these files, representing all policies issued since 1925 to policy holders with such impairments or occupations. This may seem like a rather large number of cards, but since about 400 medical impairment classes and 250 occupational classes have been studied in the last twenty years, the number of exposures and the number of deaths in some of the classes are very small.

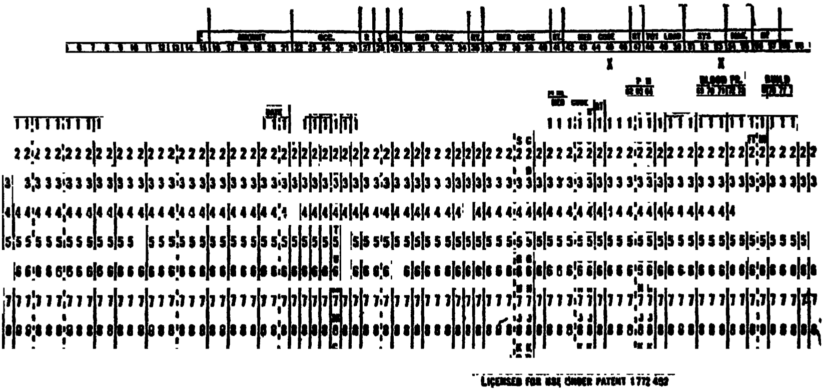


FIGURE 3
MEDICAL AND OCCUPATIONAL STUDY DETAIL CARD

These files are so constituted that data on the main medically examined standard business and on all the special classifications can be combined with such data from the larger life insurance companies to give over-all intercompany experience. This pooling of data is made, analyzed, and published by the Joint Committee on Mortality of the Actuarial Society of America and the Association of Life Insurance Medical Directors.

THE ANALYSIS OF MORTALITY AND OTHER STATISTICS

(a) *Sampling errors in death claims.* The primary problem facing the actuary is the efficient estimation of the true underlying probabilities from the statistics available. With respect to the large bulk of standard medically examined business, such large numbers are involved that sampling errors in the deaths themselves at each age and policy year are relatively small. The only serious complication arises from the fact that the primary interest is in the financial effect of claims; it is customary, therefore, to analyse the mortality rates for the most part in terms of the amounts of insurance exposed and the amounts of claims paid. However, it is a relatively simple matter to take into account the effects of variation in the amounts of insurance upon the standard deviations used to test the reliability of the statistics [1].

For instance, the standard deviation in the mortality rate by amounts, defined as the ratio of the total amount of death claims in dollars, over the total amount of insurance exposed in dollars at age x is

$$R_x q_x \sqrt{\frac{l_x}{E_x}\,\theta'_x}$$

where

$$R_x = \sum s^2 l_x^s \Big/ \sum s l_x^s$$

q_x = mortality rate by amounts $= \frac{\theta'_x}{E_x}$

s = amount of insurance in dollars on individual life (midpoint of class interval)

l_x^s = number of lives exposed with policies of s dollars

$l_x = \sum_s l_x^s$

$E_x = \sum_s s l_x^s$

θ'_x = total actual death claims in dollars.

A typical size for R_x is 2.

The process of graduating a crude series of values of the rate of mortality has the effect of reducing the variance of the individual rates of mortality, but this process is so thoroughly described by T.N.E. Greville elsewhere in this issue that mere reference is sufficient here [2].

On the other hand, for the investigation of the death rates in special mortality classifications, there are important problems involving the design of categories in which the data are analysed, the combination of data within a number of independent categories into relatively homogeneous groupings large enough to give statistically reliable results, and the determination of reliable measures of dispersion: on these problems, only slow progress is being made. We are proceeding gradually in the redesign of the classifications themselves as a result of our growing knowledge of the salient attributes of each impairment or occupation. The redesign probably can most efficiently be carried out by a working group of a doctor to supply the necessary clinical information, an underwriter to supply information concerning the acceptance of each risk, an actuary to supply information concerning the monetary effects, and a statistician to provide a proper appreciation for homogeneous and significant categorization in the light of previous statistics.

(b) *Sampling errors due to exposure sampling.* So far, the discussion has only been concerned with sampling errors in the estimation procedures arising from the limited size of the experience from death claims. In some instances, because of the cost or time involved in bringing up to date a file needed for determining exposures required in some particular estimation problem, at the present time consideration is being given to taking, say, a 10 per cent sample of all policies issued in a particular period by pulling out record cards with policy numbers ending in a particular digit. This sample will be studied for the point under investigation, and the records characterized by this point will be broken down into the necessary number of smaller relatively homogeneous categories, thus giving the exposure to risk of death in each such category. These exposures will then be combined with a 100 per cent sample of the corresponding death claims from the death claims file; the mortality rate is then the ratio of death claims to the exposure estimated from the sample, viz.,

$$q = \frac{\theta}{E}$$

where q = mortality rate
θ = deaths
E = total exposure as estimated from the sample exposure (sE)
100 s = percentage sample used in determining exposure

In the case of an investigation with large exposures in each category with lives used as units (in lieu of amount units), the variance of the mortality rate determined by such a procedure is larger than the variance in the mortality rate determined from a 100 per cent exposure sample, in approximately the ratio

$$1 + \frac{(1-s)q}{s(1-q)} \cdot$$

In the event that q involves a number of ages, and the exposure is taken over a period of years so that each life may enter into the exposure a number of times at different ages, this ratio is increased by a serial correlation effect to the approximate value

$$1 + \frac{T(1 - s)\bar{q}}{s(1 - \bar{q})}$$

where $\bar{q}$ is a type of weighted average over-all mortality rate and
T is an average period of exposure per life in the category.

The introduction of amounts of claim for the purpose of estimating the mortality rate by amounts further complicates the problem. This general problem may be found discussed in the *Transactions* of the Actuarial Society of America [3, 4].

THE EXPERIENCE RATING PROBLEM IN GROUP INSURANCE

A very important type of insurance policy today is the group policy, which insures under a single contract the health, retirement, and death benefit needs of all employees in certain employment categories of a single employer. Under this type of coverage there is a very complicated statistical estimation problem involved in determining the claim charges (credits in the case of annuities) to be assessed against each employer in the annual net cost. In order to establish proper equities and to assure relatively stable cost to the employer from year to year, it is necessary to base these claim charges partially on each employer's own experience and partially on the over-all experience in the particular premium group under which the employer's contract is classified. Naturally, the larger the number of employees and the older the employer's contract, the more weight can be given to the employer's own experience.

One type of formula used in the case of group life insurance for the claim charge against a particular employer is the following:

$$\frac{P}{A + P}K + \frac{A}{A + P}kP$$

where P = total premiums received from the employer during a stipulated period of time
K = actual claims in the employer's group during the same period
k = ratio of total death claims to total premiums over the whole group life insurance experience for the same period
A = constant chosen usually by practical considerations

This function in the case of K and P measured on a "lives" basis has been determined theoretically by simple probability considerations by R. Keffer [5]. Mr. Keffer assumes that the ratios, r, of the mortality rates for each group to the mortality rates for all groups combined are independent of age and are distributed according to a Pearson Type III curve as follows:

$$y = \frac{(m+1)e^{-(m+1)r}[(m+1)r]^m}{m!}$$

where m is a parameter related to A in the previous paragraph. Since the probability of death is small, the Poisson Law gives the probability of experiencing d deaths in a group governed by a ratio r as

$$\frac{e^{-rc}(rc)^d}{d!}$$

where c is the expected number of deaths based on the experience of all groups. Therefore, the probability of the ratio being of value r and of there being d deaths is

$$yz = \frac{(m+1)^{m+1}c^d(m+d)!}{(m+1+c)^{m+d}m!d!} \cdot \frac{e^{-(m+1+c)r}[(m+1+c)r]^{m+d}}{(m+d)!}.$$

Theoretically m could be determined by fitting this function to the company experience, although the resulting value might prove impractical. The mean value of r is then found from this to be

$$\bar{r} = \frac{m+1+d}{m+1+c}$$

so that the mean value of the deaths in the group is

$$\bar{r}c = \frac{m+1}{m+1+c}\,c + \frac{c}{m+1+c}\,d$$

which may be used as an estimate of the average number of deaths properly assignable to the group for cost purposes. If we set $m+1=Ak$, $c=kP$, $d=K$ on the assumption that each life is insured for unity, then

$$\bar{r}c = \frac{A}{A+P}\,kP + \frac{P}{A+P}\,K \text{ as set forth above.}$$

There is a need, however, for a reconsideration of the theory of this problem, taking into account distributions of exposures and deaths by age and amounts of insurance in the whole premium classification and in the individual group. Functions of a stochastic variable in the theory of risk as developed by Professor Cramér and others of the Scandinavian actuarial school may prove to be useful tools in obtaining insight into this problem [6]. It should also be mentioned that the casualty insurance actuaries have given considerable attention to this general problem [7, 8, 9].

THE CHECKING OF FILES

An interesting problem exists in connection with the checking of such files as the special mortality class files from which basic statistical data are drawn. It is desirable to have some trustworthy means of confirming the accuracy of the files without checking the whole back against ponderous original records. We are concerned with errors of various levels of importance, such as, for instance, the entries of deaths and withdrawals. We are currently examining the feasibility of setting up acceptance criteria to be applied by a binomial sequential analysis procedure [10] to a special mortality class file of some 500,000 I.B.M. punched cards.

We have in mind sorting out sequential samples in groups consisting of all cards in a particular impairment group characterized by an impairment code punched in the cards. Such a sample will involve cards into which entries of transactions should have been made over the period involved in our investigations. Each card will be compared with basic records with respect to all transactions such as, for instance, death and withdrawal, and a tabulation will be kept of the number of cards on which each type of transaction should have been made and the number of errors found in each type of transaction. In the case of death transactions, for instance, we feel that we can sustain an error of the order of 2.5 per cent of all death transactions without seriously biasing our conclusions, bearing in mind that sampling errors exist in any case by reason of the limited extent of our data. Our criteria might then consist of keeping below 2 per cent the probability of deciding the file needs complete rechecking when the actual errors are below 1 per cent and of keeping below 1 per cent the probability of accepting the file as satisfactory when the actual errors are above 5 per cent. The average sample number is below 300 for these particular criteria, so that probably no more than 600 transactions would have to be investigated before the test procedure terminated in the acceptance or rejection of the file. Depending on the percentage of death transactions among the cards sampled, which is largely a function of the period of years covered, we would hope to investigate no more than 10,000 to 20,000 cards out of the 500,000 cards. If the file does possess the desired degree of accuracy, this procedure appears to offer an economical means of confirming the fact. Of course, if too many errors are present in the file, an extensive effort is indicated to correct it.

BIBLIOGRAPHY

[1] Cody, D. D., "Standard Deviation in the Rate of Mortality by Amounts," *Transactions of the Actuarial Society of America*, vol. 42, 1941, p. 69.
[2] Greville, T. N. E., "Recent Developments in Graduation and Interpolation," *Journal of the American Statistical Association*, vol. 42, June 1948.
[3] Campbell, G. C., "A Study of the Variance of the Observed Death Rate when the Exposure is Estimated from a Sample," *Transactions of the Actuarial Society of America*, vol. 46, 1945, p. 59.
[4] Coates, R. P. and Cody, D. D., "Observations on Exposure Sampling Procedure," *Transactions of the Actuarial Society of America*, vol. 47, 1946, p. 311.
[5] Keffer, R., "An Experience Rating Formula," *Transactions of the Actuarial Society of America*, vol. 30, 1929, p. 130.
[6] Cramér, H., *On the Mathematical Theory of Risk*, Skandia Insurance Company, Stockholm, 1930.
[7] Bailey, A. L., "Sampling Theory in Casualty Insurance," *Proceedings of the Casualty Actuarial Society*, vol. 29, 1942, p. 50 and vol. 30, 1943, p. 31.
[8] Bailey, A. L., "A Generalized Theory of Credibility," *Proceedings of the Casualty Actuarial Society*, vol. 32, 1945, p. 13.
[9] Whitney, A., "Theory of Experience Rating," *Proceedings of the Casualty Actuarial Society*, vol. 4, 1917–1918, p. 274.
[10] Wald, A., *Sequential Analysis*, Wiley, 1947.

THE EDGE-MARKING METHOD OF ANALYZING DATA

L. L. Thurstone
The University of Chicago

About thirty years ago the writer introduced a simple card method of analyzing data. The method has been used by many students and others. It is especially adapted to problems with experimental populations of not more than about four hundred cases in which it is desired to analyze the relations of a large number of attributes. The number of attributes can be rather extensive, up to three hundred if necessary. The writer's 5 by 8 cards for edge-marking have been reprinted many times. Most of the users of these cards are familiar with only one or two applications. The purpose of this paper is to describe enough examples to illustrate the versatility of the edge-marking method.

Originally the provocation for devising the method was the frequent necessity for analyzing data on populations of several hundred subjects in situations where the punched card methods were not available or where they were too expensive. This was often the case with students who could not afford the expense of punched card methods and in which it was often questionable whether the size of the problem would justify the punched card procedures. When the number of attributes exceeded the number of columns on the punched card, it was necessary to resort to double punching with consequent complications in handling the analysis. Methods have been available for punching holes or slots in the edges of cards. These methods enable one to handle a large number of variables and they are often very useful. Special equipment is necessary for their use which is often not available for students who work on problems of limited size. The card method that we have been using increases still further the number of attributes that can be recorded on each card because we use pencil marks instead of holes along the edges of the card. This makes possible the use of both sides of the card for different attributes. The total edge length that is available is therefore twice the perimeter of the card. An ordinary 5 by 8 card provides 52 inches of edge length along which to record data about the individual who is represented by the card. Statical problems can be handled by a student alone without special equipment beyond the printed cards. For many years the edge-marking method has been known as the "slant-cut" card system, but the slant-cut on one corner of the card has only incidental reference to the method of statistical analysis with these cards. The cut in one corner of the card serves the same purpose as the corner-cut on the usual I.B.M.

punched cards, namely to make conspicuous any card that has been inadvertently turned over.

In Fig. 1 we have a diagram of the 5 by 8 card. Each card is first

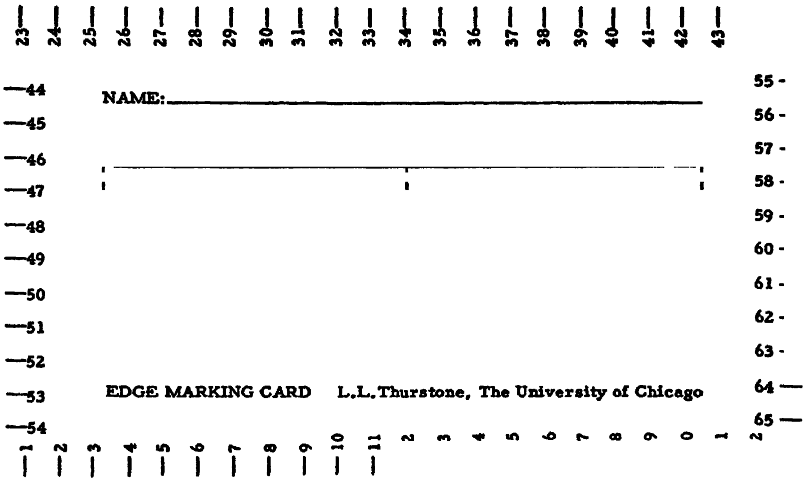


FIGURE 1

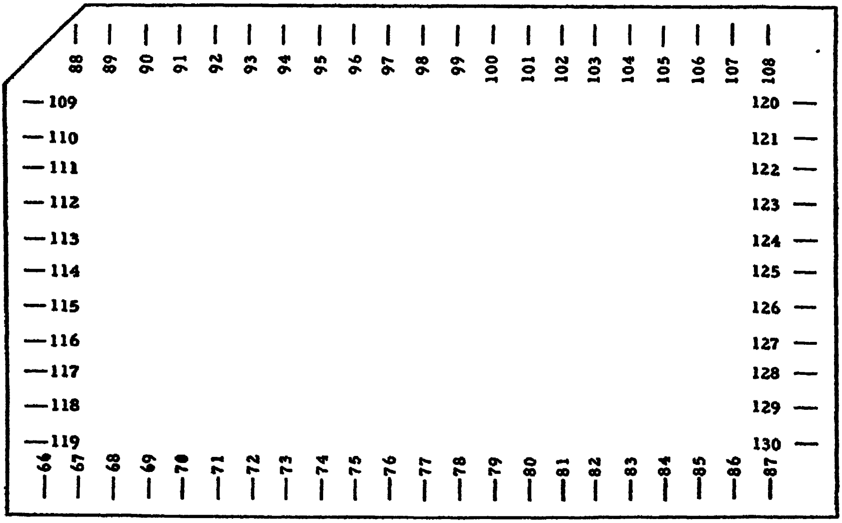
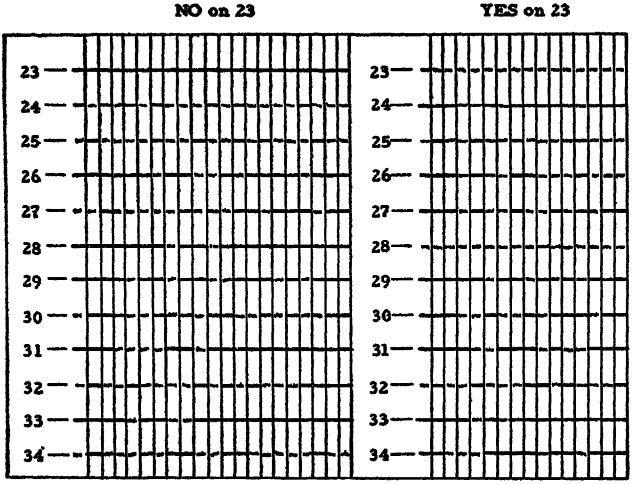


FIGURE 1A

identified with the name of the individual subject who is represented by the card. Around the edges of this card there are 130 short lines, 65 on each side of the card. Each line extends to within a short distance of the edge of the card, and each line is identified with a number. A master list is prepared showing the attribute that is represented by each code number. It is often desirable to represent the presence or absence of each attribute. This is usually done by black and red pencil marks which continue the printed line to the edge of the card. When information is lacking about an attribute, the corresponding space is left blank. Accentuation of the attribute can be represented by a double red line, or a double black line, at the appropriate space on the edge of the card. These ideas can be varied indefinitely to suit particular problems. We have found that it is not desirable to use many colors. We ordinarily limit the use of color to red and green together with black. In some problems the several gradations, or class-intervals of a variable, can be represented by several colors. In other cases, where there is plenty of room on the card and the number of variables is not large, several spaces are provided for the different gradations or class-intervals of each variable. Again, these decisions are adapted to each problem.

Perhaps the most common application of the edge-marking method is in the preparation of four-fold tables for a large number of attributes. As an example, suppose that each attribute has been given a code number and that the presence of an attribute is denoted with a red mark while the absence of the attribute is denoted with a black mark at the edge of the card. For this example we shall assume that only the code numbers 23 to 34 inclusive have been used in the positions shown at the top of the card in Fig. 1.

When a card has been prepared for each subject in the experimental population, let us start with the tabulation of all four-fold tables that involve attribute 23. The cards are stacked and then fanned so as to expose the edges as shown in Fig. 2. The cards are then separated according to the edge marks in the position 23. All cards with black marks at this position are sorted into one pile, and all cards with red marks are sorted into another pile. We shall assume here that there are no blank entries. Each of the two piles is then fanned so as to expose the edges as shown in Fig. 2. The cards that represent negative answers to item 23 will show a sold black line at position 23. The other stack of cards to the right will show a solid red line in the same position. In the figure we are using cross hatching to represent the red color. The cards with the solid red line at 23 represent the individuals with positive answers to item 23.



NO on 23		YES on 23	
NO	YES	NO	YES
21	0	0	16
7	14	12	4
6	15	4	12
15	6	6	10
10	11	8	8
19	2	1	15
15	6	12	4
11	10	10	6
16	5	11	5
9	12	6	10
18	3	13	3
7	14	11	5

	NO 24	YES 24	
YES 23	12	4	16
NO 23	7	14	21
	19	18	39

FIGURE 2

The two piles are separated slightly in stacking them so as to make the code numbers visible for both piles as shown. The figure extends only to item 34. A data sheet is prepared as shown on the right side of the figure. This table is prepared for recording the four cell frequencies for each four-fold table. The first column shows the number of black marks in each row of the first pile. The second column shows the number of red marks in each row of the first pile. The third column shows the number of black marks in each row of the second pile, and the fourth column shows the number of red marks in each row of the second pile. Consider a particular example as shown in the four-fold table at the botton of the figure. This table shows the association between the attributes 23 and 24. The association is evidently negative in that those who have attribute 23 tend to be lacking in attribute 24. This fact can also be seen by inspection of the row of edge marks at the level 24. It is immediately apparent that the proportion of black marks is larger for the second pile than for the first pile at the level 24. Since there are as many as twenty marks on a single edge of the card, it is evident that one can inspect the relations of attribute 23 with as many as twenty other attributes in one position of the cards. Twenty four-fold tables can be prepared from this position of the cards.

When the cards have been inspected, or counted, for all the attributes on one edge, the two piles can be turned to each of the other edges and the process can be continued until all of the four-fold tables have been prepared for attribute 23 which determined the separation of the cards into two piles. There are eight edges to be thus examined, four on each face of the cards. The cards can then be sorted into two piles according to each other variable that is to be analyzed.

Sometimes it is desired only to inspect the cards without counting the actual frequencies. After separating the cards into two piles according to presence or absence of a particular attribute such as a criterion measure, one may inspect the cards in search for those other attributes that correlate with the criterion measure. If the proportion of red and black marks is about the same in the two piles, one may skip the comparison without bothering to count the actual frequencies. If we do that for the example of Fig. 2, we find, for example, that attribute 26 has some relation to 23 even before counting the red and black marks. When we come to 28, we see that this attribute has a striking relation with 23 because there are many black marks on the left but only one on the right. Similarly, item 34 shows some relation to 23 because of the relatively greater density of black marks on the right. This example illustrates the use of the edge-marking method for inspectional purposes where only some of the frequencies are actually counted. In attaching significance to the relations found, it must of course be remembered that significance is determined in part by the total number of relations that are inspected.

In the present example we have assumed that there were no blank entries on the data to be recorded. If the original data contain blanks for some of the items, these can be represented by merely omitting the edge mark for the items in question. In some problems, it is of interest to know which questions draw blanks in the answers, and sometimes the blanks are considered to be significant as evasions in response to some questions. The blanks can be inspected as to relative frequency in the two groups of subjects that are being compared. When the cards are stacked as shown in Fig. 2, the relative frequencies of the blanks in the two piles can be inspected and counted if they are significantly different.

We shall list here a number of typical uses of the edge-marking method.

1. The preparation of four-fold tables for any number of pairs of variables. These can be prepared in a fraction of the time that would be required by ordinary methods of tabulation and counting.
2. The preparation of tables of multiple classification. These are the cases in which the variables are classified into more than two categories or class-intervals for the calculation of Chi square and other indices.
3. The inspection of data for questionnaires to discover significant association between variables without counting all of the frequencies.
4. The item analysis of questionnaires, psychological and educational tests, application blanks, and other schedules against some criterion or dependent variable.
5. The reduction of a large number of attributes or variables to a limited number of groups or clusters. This procedure is especially useful when the problem is to determine the dimensionality of a domain that is represented by a large number of discrete items. The composite scores for the clusters can then be studied by factorial methods.
6. The calculation of composite scores for a group of attributes. Each individual score may be the number of attributes in a group of attributes which are positive in that individual. This calculation is done with a scoring frame to be described.
7. The scoring of an individual performance when the subject classifies or sorts cards. In a factorial study of perception the writer included several tests of the optical illusions.[1] These were arranged on cards which the subject was asked to sort to the right or to the left according to the apparent magnitudes in the diagrams on the cards. The scoring was greatly simplified by edge-marking. The examiner merely picked up the stack of cards which the subject had placed in the right-hand compartment and turned them over. The number of right responses was counted by the edge marks that were on the backs of the cards so that they were not visible to the subject.

   When the examiner turned over the cards, he arranged them so as to expose the edges as shown in Fig. 2. He placed the stack of cards on a data sheet as shown at the right of the figure. A frequency distribution was then made by counting the number of edge marks in each row. This procedure gave immediately the distribution of magnitudes because each card was edge-marked on the back side at a position which designated the magnitude of the right-hand figure in the optical illusion on that card. The cards were then shuffled for the next subject.

---

[1] L. L. Thurstone, *A Factorial Study of Perception*, The University of Chicago Press, Chicago, 1944, tests 27–31, inclusive.

This work of scoring was done in a fraction of the time that would have been required by ordinary methods of tabulation. The Minnesota Multiphasic test has also been scored in this manner with edge marks on the backs of the cards. These are made by the examiner to simplify his work, but the edge marks could be printed on the cards by the manufacturer.

8. Bibliographic cards can be edge-marked to very great advantage. In every field for which a bibliography is assembled, there are a number of categories or subheadings. Each journal article or monograph is represented by a card. Instead of preparing duplicate cards for cross reference purposes, one card can be edge-marked on the back, or on the front, to show the several categories for which the article is relevant. If it is desirable to collect all the references in a particular category, one need only assemble all of the cards that have an edge mark in a particular space. When the compilation has been completed, the cards can be re-sorted into the general alphabetical file. This method saves the labor of duplicating cards or of writing explicit cross references for each card. When bibliographic cards are prepared and used by an investigator, it is often a matter of considerable importance to reduce the clerical labor to a minimum. A system that is cumbersome and which requires much writing is likely to be discontinued.
9. A practical application of the same kind can be made on a mailing list. If every person on the mailing list is represented by a separate card, it can be edge-marked to indicate the categories to which the name belongs and the kinds of material that should be mailed to that person. The fact that certain material has actually been mailed to a group of individuals can be indicated by edge-marking a stack of cards with a single pencil stroke at the appropriate space on the cards when they are fanned so as to expose all of the edges. The mailing of different types of material can be controlled from one master card file. These methods are not necessarily the most appropriate in large offices.

10. The filing of library cards and the checking of the file could be facilitated by edge-marking on the cards. The library classification could be represented by a series of marks on the edges of the cards so that the marks would be visible when the cards are stacked tight in a drawer. Any card that is misplaced would be immediately seen because its markings would not be continuous with those of adjacent cards. In this case it should not be necessary to fan the cards because the marks could be printed on the edges when the cards are stacked. This could be done by the printer when the cards are manufactured. The writer has suggested this application to librarians but the method evidently has not yet been tried.
11. If it is desired to present a set of cards in a uniform order to all subjects, the desired order can be represented by edge-marking either on the front or back of the card. When each subject has finished his categorization of the cards, they can be re-sorted by means of the edge-marking into the original order for the next subject. The alphabetized order can be represented by edge-marking because the edge-marking is perceived more quickly than the aphabetical order.
12. When paired comparison data are to be tabulated, the proportion of judgments, $j>k$, in which j is preferred to k, can be facilitated by edge-marking. This adaptation can be made in several ways depending on the original form of the tabulations.
13. A test form or preference schedule can be so arranged that the subject's responses are made at the edges of a card, the central portion of which is devoted to the text. The subject's responses can then be analyzed by the edge marks which he has made himself. This procedure saves additional time because the edge-marking is then already made by each subject.

For some types of statistical problems we have used a rack for holding the cards. This application can be illustrated by an example in which we used the larger card, namely, $10\frac{1}{2}$ by 12 inches. This was the largest size of card that could be filed in a vertical filing case of standard dimensions. In using the large size card it is best to have a rack on which to place the cards for the kind of comparison that was described with Fig. 2. The rack consists of a board that is mounted at an angle of about sixty degrees with the surface of the table. At the lower edge of the board is a ledge on which the cards rest. Two or more groups of cards can be arranged on this board for inspection of the edge marks. When using the rack for counting frequencies, especially on the large cards, the work is facilitated by using a long flexible ruler or preferably a horizontal rule that can be clamped in position for each row when the frequencies are being counted or inspected.

The use of the scoring frame will be described in terms of a recent study in the Psychometric Laboratory at The University of Chicago. A schedule of 330 questions had been assembled on temperamental characteristics. The subject was asked to respond to each question with one of three responses, namely, yes, no, or a question mark. The 330 responses were edge-marked on a single card for each subject. The assumption was made that the large collection of questions represented a much smaller number of factors, dimensions, or types. The purpose of the study was to make a factorial investigation of the dimensionality of the domain that was represented by the entire list of questions. For this purpose, it was desirable to discover first the main groupings of the items and to make the grouping in such a manner that all of the items in each group should be as closely related as possible.

The items were first arranged in tentative groups according to preliminary interpretations. Then we wanted to assign to each subject a composite score for each group of items. For example, one group of items was concerned with rate and amount of activity. A composite score for such a group of items would be the subject's self-rating on general activity. A very active person would be expected to have a high composite score while a calm and deliberate or phlegmatic person would be expected to have a lower rating on this characteristic. Such a composite rating does not imply that the trait will be revealed as a primary factor. It might turn out to be a combination of other more general or fundamental factors. Before such a composite rating can be accepted even for preliminary study, we should have some assurance that the several individual items in each group are highly correlated with the composite score for the group. It was for this type of analysis that we wanted a composite score for each subject in each group. Those items which did not show close relation with the composite score were to be eliminated from the group and other items would be added by inspection of the edge marks.

The composite scoring frame is represented diagrammatically in Fig. 3. For simplicity of explanation, we shall assume that all of the items in the group to be analyzed are represented on the upper and right-hand edges of the front face of the edge-marking card. The scoring frame has the four sides A which enclose a rectangular space equal to that of the edge-marking card. One of the individual cards is placed into this rectangular space. The upper and right-hand edges of this card are shown at *B* in Fig. 3. A master card *C* is prepared for each group of items. This card is slightly smaller than the edge-marking card. The master card in the figure has been moved in the direction *D*, toward the lower left corner of the frame. When the master card is moved toward the lower left corner of the frame, the edge marks become visible along the upper and right-hand edges as shown in the two edges *B*. Let the cross hatched marks represent red edge marks on the individual card.

Let us suppose that nine items have been grouped in the belief that they are closely related. These nine items are represented by edge marks on the master card *C*. Red and black edge marks have been made on the master card according to whether positive or negative responses to the items represent presence or absence of the postulated general trait. When the master card is in the position shown in Fig. 3, one can readily count the number of items on which the subject responded in the expected manner if he excelled in the group trait that is represented by the master card. The composite score for the individual subject is the number of items on which he responded in a manner similar to that which is marked on the master card. In the example of Fig. 3 we find that the individual subject agreed with the master card on seven of the nine questions. On the top edge, for example, we find that the subject

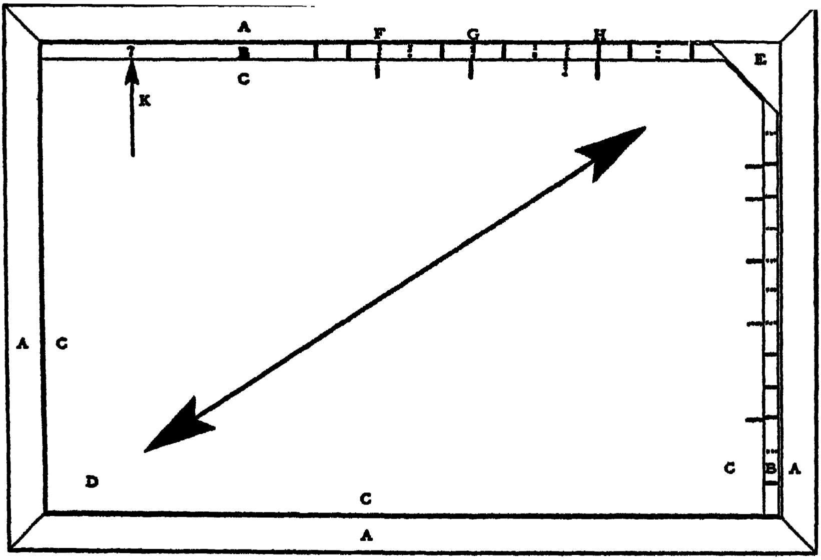


FIGURE 3

made two positive responses and one negative response that agreed with the master card. The responses at *F* and *H* agree with the master card while the response at *G* disagrees with the master card. The slant-cut at *E* is also made on the master card so as to insure that the master card and the individual card are properly matched.

When the count for the composite score has been made, the result is recorded on the individual card. The arrow *K* on the master card indicates where the composite score for this group of items should be recorded on the individual card. The composite score 7 is now recorded on the individual card. The same procedure can be followed for composite scores on other groups of questions or items. The composite scores for different groups are, of course, recorded in different positions on the individual card. When all of the composite scores have been recorded on the edge-marking cards, these can be tabulated in a frequency distribution. The median composite score is then determined. The cards can then be marked with red and black to indicate whether the composite score for an individual is above or below the median or mean. The cards are then ready for the preparation of four-fold tables to show the relation between the individual item responses and the composite score for each group. Similar analyses can be made for quartiles or for any other groups. Those items that do not agree with the composite can be eliminated and other items can be discovered by inspection that should be included instead.

In the present example, we have assumed that all of the items in a group were on the two edges of the card which are simultaneously visible. If the items spread over more than the two edges, the master card is moved in the direction *E*. That exposes the edge marks along the left edge and the bottom edge of the card. If the group items extend over both faces of the card, then the individual card and the master card are turned over so that the counting can be continued on the second face of the card in the same manner. The master card is then also marked on both faces. These procedures are much easier to show with actual manipulation than with the awkwardness of verbal description. Item analyses of various kinds can be quite easily handled with these edge-marking methods.

The edge-marking cards can be marked directly from the original data. When the data are arranged in columns, it is convenient to place the edge of the card next to a column of data and to transfer the column of records directly to the adjacent edge of the card. In order for this procedure to be feasible, the spacing of the columns on the original data sheets should be the same as that of the edge-marks on the card.

One of the writer's students, Mr. Albert Hunsicker, made a practical application by using blank cards instead of the printed cards. He adopted a spacing for the edge-marking that matched the spacing on his data sheets. The writer has used blank cards in the same manner. The card shown in Fig. 1 was arranged to fit double spaced typewriting. If the original data sheets have such spacing, then the data can be transferred to the edge-marking card directly by placing the edge of the printed card next to the column of data. This method of making the edge marks saves much time and avoids clerical errors.

In making the edge marks it is advisable to avoid a soft pencil because such pencil marks are likely to rub off to adjacent cards when they are handled. We have used red and black India ink to advantage on cards which are to be handled for a large number of tabulations. A rather fine pen point seems to be adequate for edge-marking. If the cards were to be manufactured in large quantities, they might be made with a border that is slightly thinner than the rest of the card. The edge marks would be made on this slightly thinner border and they would be less likely to rub off onto adjacent cards even with extensive handling. The difficulty is avoided by using a medium-hard pencil or by using ink.

The flexibility and economy of the edge-marking method of classifying and analyzing data justify more general use than it has had in the past. The method is especially useful for those problems in which the investigator works either alone or with some clerical assistance and where tabulating equipment for punched card methods is not readily available. For some problems with relatively small populations and with many attributes for each individual subject, the writer prefers the edge-marking method to the punched card procedures because the significant relations can be identified by inspection without actual counting and tabulation. The several examples of this paper are intended to illustrate the versatility of edge-marking which can be adapted to a great variety of purposes.

THE WAR PRODUCTION BOARD'S STATISTICAL REPORTING EXPERIENCE, II and III

DAVID NOVICK
University of Puerto Rico
AND
GEORGE A. STEINER
University of Illinois

PART II

CHARACTERISTICS OF THE WPB REPORTING STRUCTURE

> **An understanding of the scope of the WPB reporting structure demands more than a knowledge of its development as described in the first article of this series. Further insight into its intricacy and breadth is given by an analysis of its quantitative and qualitative characteristics.**

NUMBER OF WPB QUESTIONNAIRES

NUMERICAL statements concerning WPB questionnaires must be viewed with caution. The numbers of requests for data, and the figures purporting to reveal their character, have serious limitations as units of measure. Reporting forms varied in length and complexity from a multipage form, such as PD-25A, to a simple registration statement requesting only a name and an address. Some complicated forms were often little used by management and simple questionnaires frequently had a high administrative value. Such factors should be kept in mind in appraising the following analysis.

From March 1941 to September 1945, a total of 4,083 different forms were approved for issuance by the WPB. This figure includes a total of 2,397 repetitive or recurring forms, and 1,686 "one-time" forms. These figures do not take into account the "wild-cat" forms submitted to industry without formal approval, those forms continued by other governmental agencies and trade associations for use by the WPB, forms submitted for formal approval but disapproved (the disapproval rate was about 10 per cent of the forms submitted), or the unknown but substantial number of forms that were discussed and then withdrawn by the originating unit before formal submission. The total figures compare with 309 questionnaires issued by the War Industries Board in World War I.

QUANTITATIVE ANALYSIS OF REPETITIVE FORMS IN USE IN DECEMBER 1943

A quantitative analysis of the recurring questionnaires in existence at one point of time gives an enlightening cross-section view of the WPB reporting structure. The study reveals the character of data used by the WPB management since continuous surveys were, for the most part, the instruments upon which administrative action was founded. The techniques used on forms at one point of time, however, should not be considered representative of the entire war period. Certainly the discussion in the previous article indicated the hazards in drawing such a conclusion.

The following study of questionnaires in use in December 1943 was derived from an *Analysis of the Use of WPB Forms* prepared by the General Statistics Staff of the WPB and published in December 1943.[1] During December 1943 there were in use a total of 707 recurring forms. It is useful to compare this total to counts of repetitive reports in use at various points of time. The numbers were: September 1942, 609; January 1943, 586; May 1944, 903; and August 1945, 442.

In the following study a number of one-time and irregular reports are added to the 707 repetitive questionnaires in use in December 1943 to make a total of 786 forms analyzed. The 786 forms are classified into the following uses: statistical, 508; application, 305; scheduling, 48; and other, 95. Included in the "other" category are registrations, certificates, bills of material, reports of surplus stock, and appeals. Since one form often had several uses the total number of these 786 forms amounted to 956.

WPB questionnaires and reporting instruments were predominantly of three types: statistical, application, and scheduling. Statistical questionnaires were those designed to obtain information for use in policy-making and in serving as guides to action. They were not vehicles for administrative action. Applications were primarily designed to serve as instruments for authorization by the WPB to industry to take certain action. Most of these applications required the submission of data that could be used for statistical purposes. The WPB learned that reporting was facilitated on a *quid pro quo* basis, that is, when the submission of data resulted in authority to industry to do something. Scheduling forms were those designed primarily to obtain information on the dates at which a manufacturer proposed to ship a designated item to a specifically identified customer. On the basis of such reports the WPB "froze" or modified schedules. Many questionnaires embodied all three of these characteristics.

---

[1] See also *Journal of The American Statistical Association*, Vol. 39, No. 226, 1944, pp. 144–154; "Wartime Facts and Peacetime Needs," by Vergil D. Reed.

The type of data requested on statistical forms is indicative of the sort of information which the WPB felt it needed for operations. The following summary derived from an analysis of the 508 forms used for statistical purposes, shows the number of forms requesting the type of information specified:

Capacity	70
Production:	
Actual	204
Scheduled	111
Shipments	271
Consumption	181
Receipts or accepted deliveries	178
Inventory	339
Orders:	
Unfilled	147
New	119
Cancelled	52
Requirements	102
End Use Pattern	128
Other	151

The frequency of return of all 786 forms and the 508 statistical forms is summarized as follows:

Frequency of Return	All Forms	Statistical Forms
One-time	79	42
Daily	1	0
Weekly	3	2
Semi-monthly	8	6
Monthly	404	351
Bi-monthly	3	0
Quarterly	88	71
Semi-annual	1	0
Annual	1	1
Irregular	198	45

The proportion of the total maximum number of respondents filing each form was important for statistical reports. The extent of coverage of the 508 statistical forms was as follows: complete coverage, 383; large firms only, 55; sample of large, medium, and small firms, 22; and coverage unknown, 48.

Analysis of the type of respondents shows some of the broad areas covered by the WPB's reporting. Of the 786 forms, 430 were filed by manufacturers; 88 by other types of processors, such as converters, tanners, refiners, foundries, and mills; and 159 by "consumers" who were largely, but not entirely, manufacturers filing forms as consumers of raw materials. At the distribution level there were 121 forms filed by other classes of respondents such as warehouses, distributors, wholesalers, importers and retailers. There were 6 surveys the respondents of which were utilities; 15 covered government agencies, and 85 covered other groups. (The total of groups is 904 and results from the fact that one form often covered several groups.)

The number of respondents filing WPB questionnaires varied from one to hundreds of thousands. Form PD-789 (producers monthly report of fibrous glass production), for example, was filed by one company. Preference-rating certificates, such as PD-1A, and various other application forms were filed by thousands of individuals, plants, and companies. Contrary to popular impressions the average number of respondents to most reporting forms was relatively small. The following summary table shows that the number of respondents was fewer than 100 for more than half of all statistical forms. An astonishing number of forms was collected from 25 or fewer respondents. Only 54, or a little more than 1 per cent of the total, were submitted by more than 1,000 respondents.

TABLE I

	Number of Respondents Classified by Reporting Periods				
Number of Respondents	One Time	Monthly	Quarterly	Other	Total
N.A.	5	10	4	8	27
10 and less	1	31	4	7	43
11–25	4	46	5	8	63
26–50	2	53	6	8	69
51–100	4	60	13	6	83
101–200	3	48	10	5	66
201–500	6	48	10	3	67
501–5000	14	39	15	9	77
over 5000	3	8	2	0	13
TOTAL	42	343	69	54	508

Most of the questionnaires were related to a WPB order, such as an L or M order. Of the 786 forms, 178 were not directly related to orders; the remainder, 608 forms, were related to more than 300 different orders.

WPB REPORTS FILED BY A MEDIUM-SIZED PLANT

It is informative to study the quantitative characteristics of the WPB reporting system by analyzing the reports filed by a single industrial plant.* For this purpose a medium-sized plant which employed approximately 5,500 persons and produced electric motors, various types of electrical equipment and miscellaneous war products was selected for study. Reporting requirements for various products were so diverse, however, that no one plant should be considered representative of all plants.

A total of 57 reporting forms were chosen. Excluded from the study are comparatively insignificant and irregular reports, telephonic requests for data, informal letters requesting information, various authorizing forms which were not submitted directly by the plant, and, of course, reports filed with the Aircraft Scheduling Unit, Board of Economic Warfare, War Department, Navy Department, and other such procurement agencies, whether the form did or did not carry a WPB form number. Inclusion of these items would, of course, result in a much greater total number of questionnaires and surveys actually submitted by this plant.

Table II shows the frequency throughout the war period with which selected categories of information were requested on the 57 forms. In December 1941 this plant was required to submit only 4 forms. These were statistical-type reports filed monthly; three covered aluminum data and one asked for information about copper. By the first half of 1943 the number had risen to 37 and thereafter fell off only slightly. Throughout the war period inventory data were most frequently requested while data pertaining to shipments and production were next in importance.

Neither these figures nor others in the tables reveal fully the volume and detail of information asked on individual forms. Modifications of old questionnaires and new schedules tended to request more penetrating information, in greater volume and detail, as production and distribution problems became intensified in the period from early 1941 through 1943. Thereafter there was a general tendency in the other direction, but it was not strong until the spring of 1945.

Table III classifies the type of information reported by the frequency of the reporting period. This table demonstrates clearly that when considering the war period as a whole most information, particularly detailed information, was reported on a monthly basis.

---

* The authors wish to acknowledge that Mrs. Elisabeth Joyce undertook, under the supervision of the authors, the analysis upon which this section is based.

TABLE II

Selected Categories of Information Requested on Fifty-Seven WPB Forms Filed by One Plant in Selected Periods of Time

Informational Category	Total Forms Dupl.	1941	Periods of Time							
			1942		1943		1944		1945	
			Jan. June	July Dec.	Jan. June	July Dec.	Jan. June	July Dec.	Jan. June	July Dec.
Capacity	7		2	4	4	2	2	1	1	
Production										
Actual	16	3	6	7	11	10	9	10	10	
Scheduled	10				7	8	8	10	10	
Shipments										
Dollar	8	1	3	2	5	2	3	3	3	
Unit	23	4	7	10	17	14	12	12	12	
Material Consumption										
Total	13	1	4	7	11	9	6	6	6	
By products	4	1	3	3	3	2	2	2	2	
Anticipated	7	1	1	2	7	4	3	3	3	
Receipts or Accepted Deliveries	15	1	5	8	13	10	6	6	6	
Inventory	26	3	10	12	21	19	15	15	15	
Orders										
Unfilled	12	1	1	2	6	6	7	7	7	
New	6			2	5	4	3	3	3	
Canceled	3			1	3	3	3	2	2	
Material Requirements	14	1	5	7	12	8	7	7	7	
Labor										
Use	5		3	2	3	3	3	2	2	
Requirements	1				1	1	1	1	1	
End Use Pattern	13	1	1	6	8	8	5	5	5	
Number of Unduplicated Forms	57	4	15	20	37	36	32	31	30	

The frequency with which various types of information were reported over the war period on types of questionnaires is shown in Table IV. This table shows that most information was reported on statistical-type forms with shipment, production, inventory and material consumption data predominating.

Table V on page 471 shows that most of the reports were very limited in scope. Throughout the war period the plant studied was requested to file the great bulk of its reports for individual products.

TABLE III

Selected Categories of Information Requested on Fifty-Seven WPB Forms Filed by One Plant by Time Frequency of Reporting

Informational Category	Total Forms Duplicated	Frequency			
		Monthly	Quarterly	One-Time	Irregular
Capacity	7	4	—	2	1
Production	22	13	6	1	2
Actual	16	13	2	1	—
Scheduled	10	4	4	—	2
Shipments	27	16	8	3	—
Dollar	8	1	6	1	—
Unit	23	16	4	3	—
Materials					
Consumption	18	10	5	1	2
Total	13	9	1	1	2
By products	4	1	3	—	—
Anticipated	7	4	2	1	—
Receipts or Accepted Deliveries	15	11	2	1	1
Inventory	26	14	7	4	1
Orders	15	10	4	1	—
Unfilled	12	7	4	1	—
New	6	6	—	—	—
Canceled	3	3	—	—	—
Material Requirements	14	7	4	—	3
Labor	2	—	—	—	2
Use	5	1	1	1	2
Requirements	1	—	—	—	1
End Use Pattern	13	8	1	1	3
Number of Unduplicated Forms	57	25	10	7	15

Most of the forms for which the company was held responsible were monthly summaries. The frequency with which reports were filed is summarized in Table VI on page 471.

Conclusions should be drawn from these quantitative data only after examination of other aspects of the reporting problem. These studies become more meaningful when considered in the light of outstanding qualitative characteristics of the WPB reporting system.

TABLE IV

Selected Categories of Information Requested on Fifty-Seven WPB Forms Filed by One Plant by Character of Use

Informational Category	Total Forms Duplicated	Statistical	Statistical and Application	Application	Other
Capacity	7	5	0	0	2
Production	22	14	2	6	0
Actual	16	14	2	0	0
Scheduled	10	4	0	6	0
Shipments	28	20	5	2	1
Dollar	8	4	2	2	0
Unit	23	18	4	0	1
Material Consumption	18	8	8	2	0
Totals	13	7	6	0	0
By products	4	2	2	0	0
Anticipated	7	3	2	2	0
Receipts or Accepted Deliveries	15	8	7	0	0
Inventory	26	15	7	1	3
Orders	13	8	1	3	1
Unfilled	12	8	0	3	1
New	6	5	1	0	0
Canceled	3	3	0	0	0
Material Requirements	14	2	9	3	0
Labor	5	3	0	1	1
Use	4	2	0	1	1
Requirements	1	0	0	1	0
End Use Pattern	13	4	3	4	2
Number of Unduplicated Forms	57	27	9	13	8

GENERAL QUALITATIVE CHARACTERISTICS OF THE WPB REPORTING STRUCTURE

The WPB reporting framework developed various broad qualitative patterns which gave it a distinct character. The reporting structure in its formative years, and to a certain extent in its entire lifetime, was not flexible enough to meet new fundamental problems as they appeared. It did not have sufficient substance to permit the detection of crises before they arose. Except in connection with isolated product control problems the reporting structure followed rather than preceded

TABLE V

Distribution by Time Periods of Company, Plant, and Product Reports Filed by One Plant on Fifty-Seven Forms

Time Period	Company Report	Plant Report	Product Report
Total (unduplicated)	3	2	52
1941	0	0	4
1942			
First Half	2	1	12
Second Half	0	2	18
1943			
First Half	1	2	34
Second Half	1	2	33
1944			
First Half	1	2	29
Second Half	1	1	29
1945			
First Half	1	1	29

the changing character of needed controls. The reporting structure, therefore, was molded largely by improvisations generated out of the pressures of events. Its ability to prepare for and, by so doing, to choke off serious production problems was not one of its sources of strength.

TABLE VI

Distribution by Time Periods of Irregular, One-Time, Monthly, and Quarterly Reports Filed by One Plant on Fifty-Seven WPB Forms

Time Period	Irregular	One-Time	Monthly	Quarterly
Total	15	7	25	10
1941	0	0	3	1
1942				
First Half	2	2	7	4
Second Half	4	2	11	3
1943				
First Half	10	2	22	3
Second Half	13	2	17	4
1944				
First Half	14	0	12	6
Second Half	14	0	12	6
1945				
First Half	13	0	12	5

The reporting system penetrated almost every conceivable aspect of industrial production activity, from the accumulation of total production data to cumbering incidentals. Moreover, WPB surveys did not stop with industry: they also touched farmers, housewives, retailers, wholesalers, financial institutions, and foreign Governments. The range of information sought was as expansive as the breadth of controls required to channel the productive efforts of the industrial machine. Data requirements ranged from questions on operations of giant corporations to the needs of an individual farmer who wanted to buy an alloy steel bolt to fix his tractor. There were actually millions of applications transmitted to the WPB by individuals and small businessmen to purchase items on restricted lists.

Chronologically, the main evolution of the reporting structure pertained to individual products and materials. Not only did this pattern fit the general philosophy of control but it fitted also the organizational structure of control agencies. Throughout the war period the basic fabric of the reporting structure was patterned after individual product reporting. The superimposition of comprehensive surveys upon these product reports created many intricate and serious technical problems. Except in a few areas, coordination of comprehensive surveys and individual product reports was never fully accomplished.

The most commonly understood characteristics of the reporting structure probably were the overlapping of questionnaires, inconsistency in reporting requirements, and duplication of incoming information. In the chaotic period in mid-1942, major overlapping and duplication was inherent in such broad and basic reports as PD-275, PD-25A, WPB-732, and in individual allocation schedules. Within each of the individual allocation systems the same type of duplication existed. For molybdenum, for example, a steel producer might report molybdenum requirements on his alloy melt schedule, PD-391; his consumption and stocks of molybdenum on WPB-106; detailed molybdenum requirements for melting on PD-258; molybdenum consumption on PD-259; and present his application and obtain his allocation for molybdenum, pursuant to order M-110, on PD-260. Many comparable situations arose from administrative expediency which did not admit sufficient time to eliminate such conditions through coordination.

Throughout its existence, although in lessening degree, the reporting structure revealed an amazing lack of procedural unification and standardization. This extended from questionnaire format through coverage, tabulation, and informational usage. The WPB, for example, constantly struggled with the problem of standardizing application forms relating to L and M orders. In mid-1942 the Chemicals Division devised two standard forms for its orders (PD-600 for consumers of chemicals and PD-601 for producers of chemicals) which eventually were used to implement 117 orders. This type of experience, however, was not common in the WPB and most orders carried their own specialized forms.

Technical methodology developed in the reporting structure generally evolved through three stages. The first stage was characterized by difficulties in grasping problems which reports could solve, in recognizing the applicability of old reporting procedures, or in developing new procedures to solve them. In this phase, fumbling, experimentation and false starts predominated. Out of these experiences a second stage emerged in which practical methods were sorted from impractical ones and incorporated in recognized reporting standards. The third stage was characterized by an application of these methods to the problems at hand and the development of modifications by which they were made to fit better the needs of management.

Although most methods developed along these lines it must not be inferred that there was a free interchange and free use of methods among the WPB's personnel so that tested techniques used by others were readily adopted in all areas where they were applicable. WPB personnel groups showed a strong resistance to accepting methods and procedures adopted by others. Each new problem was considered to be highly specialized and to demand highly specialized methods for its solution. Although on the surface new problems often appeared to be different from others which had arisen, most problems displayed comparable fundamental characteristics. Their solution, in so far as that could be accomplished by statistics, thus merited review of techniques adopted in other areas. But such review was never sufficiently general to result in free interchange and mutual acceptance of common experience.

The WPB reporting system was considered a "step-child" by far too many administrators. Its behavior, except by those fostering it, was often viewed with suspicion. It was characteristic of WPB personnel to disparage the results of all surveys except their own, particularly when the data from each bore on a common problem. This attitude was probably more indicative of ignorance regarding statistical methods and procedures than of unjustified affection for one's own data accumulation. Each person sponsoring and using information generally was aware of its strengths and weaknesses. But the complexity surrounding the collection and use of each set of data was such that generally nothing short of close familiarity with all details of the data could provide full comprehension of them.

The reporting structure was subject to rapid changes not only in complete basic control systems but also form details. We have already seen the outlines of swift changes in broad control mechanisms. Even quicker changes took place with particular forms. Many of the one-time questionnaires appeared suddenly and without previous notice. Among the periodic or recurring reports rapid and unannounced data changes often took place. These changes were unusually swift and substantial early in the war period but even toward its end sudden and radical alterations of old forms were frequent. In 1944 central review action was taken on a total of 2,841 questionnaires. Of this number 1,933 were revisions or extensions of existing reports. Some revisions were of no consequence; others were so extensive as to constitute a completely new questionnaire.

The reporting structure, as an instrument or tool of industrial management control, reflected either directly or indirectly all the intricate wartime production control problems. It was the impact of these problems on the use of reporting for management purposes, and the resolution, irresolution, and compromise exercised in solving them, that gave the reporting structure its character and direction. An appraisal, therefore, of the WPB reporting structure in terms of statistical reporting ideals cannot be realistic without a recognition of these influences.

PART III

CONTROLLING THE ISSUANCE OF QUESTIONNAIRES

Fortunately, in the earliest days of the NDAC, a central clearance function was established to approve or disapprove issuance of questionnaires. In retrospect, the need was obvious for a centralized review of proposed data collections to avoid duplication, to coordinate reporting, to maintain technical standards, and to insure administrative efficiency in the use of collected data. Through central control, all problems relating to questionnaires were brought to a focus.

Before a questionnaire could be approved for issuance, it was necessary to compromise a variety of pressures relating to its issuance. The range of WPB data requirements and methods was so broad and complex, the impact upon industry so difficult to measure, and the relationship between administrative responsibility and statistical procedural standards so intricate, that questionnaire issuance created thorny problems which by their very nature could not be answered to the satisfaction of everyone concerned. On the one hand, for example, were the desperate appeals for clearance of requests "without which the WPB cannot operate." On the other hand were complaints from industry that "we are producing weapons not statistics." Compromising such conflicts was difficult.

LEGAL BASIS OF CENTRAL CONTROL OVER FORM ISSUANCE

The wartime scramble for facts by Government agencies led to a substantial volume of Government questionnaires to industry. Widespread criticism of poorly devised report forms, duplication of data requests, unintelligible instructions, and the rapidity with which requests multiplied, reached a peak early in 1942. Action was not slow in coming. Corrective measures were forced on various Federal agencies, committees formed to study the problem, and various business advisory groups were pressed into service. The most visible result of these actions was the passage in December 1942 of the Federal Reports Act of 1942.

The Federal Reports Act formalized responsibilities of Federal agencies in their data collection work and provided a basis for the Bureau of the Budget to control all data-collecting activities of Government. The act specifically directed the Bureau of the Budget: "to investigate the needs of the various agencies for information from business enterprises and other agencies; to investigate the methods used in obtaining such information; and to coordinate information-collecting services in order to reduce the cost of the Government and minimize the burden upon business and industry and utilize, so far as practicable, files of available information and existing facilities of the established Federal departments and independent agencies."

The NDAC in large measure had anticipated the Federal Reports Act of 1942 and by so doing had prevented a far more chaotic situation than that which existed in the early war period. At the first meeting of the Advisory Commission to the Council of National Defense, held on June 12, 1940, it was stated that "experience had already demonstrated the necessity for establishing a central office through which all requests of the Commission for information would be cleared, that a plan was being prepared, and that a tentative suggestion as to the organization of such an office would be presented to the Commission not later than Saturday, June 15, 1940." In the summer of 1940, before the NDAC began to collect data from its own questionnaires, the Division of Statistics of NDAC, jointly with the Bureau of the Budget, studied the reporting experience of the War Industries Board in World War I. Analysis of problems faced by the War Industries Board (which did not have formal centralized form control) led to the recommendation that the NDAC should create a central clearance point for data requests originating within the Commission. A control office was created but it operated informally and had no power to compel compliance with its decisions.

Results achieved by this office in coordinating data requests and in improving the character of questionnaires issued, however, led to more formal authority. This was established, after the OPM absorbed NDAC, by the issuance of OPM Administrative Order No. 4.

Administrative Order No. 4 required every office of the OPM to obtain the approval of the Division of Statistics before any formal inquiry could be sent to industry. WPB General Administrative Order No. 6, issued in February 1942, carried this clearance requirement through the new WPB. Criteria set forth in the order to govern the review of questionnaires are worthy of note. They were:

"1. That the information to be requested is needed at the time it is to be filed and that the need justifies the effort and expense on the part of both industry and government in order to obtain it.
2. That the data are not already being secured by any other federal agency, or, if being secured, are not and cannot be made available to the War Production Board.
3. That the request includes clear and specific definitions of the required data.
4. That the required report is adapted as closely as possible to the types of records ordinarily maintained by business concerns or available to them.
5. That staff and equipment are available to tabulate or otherwise process the returns when received."

The Office of Survey Standards was established in the Division of Statistics to carry out the provisions of this Order. Passage of the Federal Reports Act of 1942 necessitated only one revision in WPB procedure: each report form in use was given an approved number assigned by the Bureau of the Budget.

There was continued in the OPM, from the NDAC, an Office of Procedures which also was responsible for the review of questionnaires. The Office of Procedures carried broad responsibility for reviewing all administrative procedures relating to all plans, programs, orders, and regulations proposed for issuance by the OPM. This Office carried responsibilities in administrative procedures comparable with those of the Office of Survey Standards in questionnaire methods and standards. As part of its operations the Office of Procedures reviewed questionnaires in terms of administrative practicability, relationships with control programs, and other procedural matters. In exercising this function there was a clear demarcation from functions of the Office of Survey Standards. The functions of both, however, shaded into each other.

This overlapping proved to be a source of strength rather than weakness. Nevertheless demarcation was required. In April 1943 a working agreement was reached to formalize previous informal arrangements. It was agreed that the Office of Survey Standards would confine its review of data requests to "statistical problems, such as: (1) review of the kind of data to be secured; (2) the form in which such data was to be secured, e.g., pounds, feet, tons, gallons; and (3) the instructions for filling in the requested data." On the other hand, the Office of Procedures confined its review to "administrative impact," that is: (1) the administrative necessity for the form, e.g., if the method of control provided in the order was questioned, the form implementing such control *ipso facto* was questionable; (2) the possibility of consolidation of forms; and (3) instructions covering the number of copies of forms which were to be filed, with whom filed, and so on. Broadly, the former matter was statistical and the latter procedural.

This agreement crystallized a procedure which had existed long before. Data requests followed a simple path. First, the Office of Survey Standards reviewed new questionnaires. If the questionnaires were approved they moved to the Office of Procedures for review against existing and efficient administrative procedures. Since our interest is primarily statistical rather than procedural the remainder, although not all, of this section concerns operations of the Office of Survey Standards. The concept of "centralized questionnaire control" does, however, include operations of both groups.

CLEARANCE PROCEDURE

Rudimentary principles of efficient action for clearance of data requests were so well recognized by those familiar with the problems involved that no serious question of clearance procedure arose. Varying circumstances surrounding the adoption of each new reporting proposal, however, often resulted in short-cuts to established lines of action. Nonetheless, most inquiries went through much the following type of critical review.

The bulk of data requests (over 90 per cent) originated in the WPB

Industry and Material Divisions. The divisions were staffed for the most part with businessmen fresh from industry, who were presumed to have some knowledge of the data needed for control and of the ability of their industry to supply the requested information. These businessmen had help from professional statisticans assigned to each division, who were presumed to have technical skill in collecting and using data. The idea that this first screening would result in a minimum of ill-conceived questionnaires, however, was not supported by the facts.

After these preliminary reviews the data request was submitted to the Office of Survey Standards. The transmittal folder contained a transmittal letter; a memorandum of justification; copies of the proposed questionnaire; and other necessary papers such as processing directions, printing requisitions, mailing instructions, and so forth. The Office of Survey Standards was required to approve or disapprove the data request within three days of submission, a time schedule not always met, or to notify the initiating unit of objections. Copies of the form, memorandum of justification, and other papers were routed to the Bureau of the Budget during this three-day period to notify the Bureau of the proposal and to allow concurrent review by the two Offices. If it became necessary to consult the unit originating the report, a representative of the Bureau of the Budget was invited to be present to prevent unnecessary repetition of interviews.

When a data request was approved by the Office of Survey Standards, a formal request for approval was made to the Bureau of the Budget under the provisions of the Federal Reports Act of 1942. Because of the practice of concurrent review, there were probably not more than a half dozen cases in which the Bureau of the Budget disapproved a form after formal request for approval by the Office of Survey Standards. Decisions to modify or disapprove data requests were discussed with the originating unit to reach agreement on issues involved. In most instances this practice led to compromises acceptable to the parties concerned. But if no compromise agreement could be reached, and where the originating unit could not be convinced of the need for withdrawal of the request, no alternative was left but to set forth formally the reasons for rejection.

Having obtained approval by the Bureau of the Budget for a questionnaire, the Office of Survey Standards affixed the Budget Bureau approval number, signed the print requisition, and forwarded the final draft of the form and the printing and mailing requisitions to the Office of Procedures. That Office, which had charge of coordinating procedures used by various divisions, was consulted in advance by the Office of Survey Standards if a proposed data request implied a change in established procedure. Consequently, obtaining written approval of the Office of Procedure was a fairly simple formality. The Office of Procedures, after review, routed the documents to the Printing and Distribution Control Division which had jurisdiction over the layout, printing and bulk distribution of all WPB forms.

This general review procedure altered little during the war period. This was partly due to legal restrictions, partly to a general recognition of the soundness of the procedures, and partly to the fact that the procedures were flexible enough to operate effectively in a rapidly shifting reporting structure. From time to time, however, this procedure was supplemented by the work of special committees.

COMMITTEE FOR THE REVIEW OF DATA REQUESTS FROM INDUSTRY

At no time during the war period was there an absence of strong industry criticism of WPB forms. WPB data requests constituted a burden on industrial operations. One of the first efforts to lighten the burden was begun in June 1942 by a Committee for the Review of Data Requests from Industry.

This committee, composed of representatives from business, the Army, the Navy, and the Bureau of the Budget had wide powers to correct shortcomings of WPB questionnaires. The establishment of the Committee, and its system of sub-committees and industry advisors, did not conflict with past clearance procedure. Rather, it created a "Court of Appeals" which constituted a new clearance hurdle.

The Committee had the support of all responsible executives in the WPB and all others interested in improving the elaborate and extensive questionnaire structure created by the WPB. Each WPB division responsible for data requests was asked by the Committee to review its questionnaires and to supply the Committee with a statement of repetitive report forms in use showing: those which should be continued, those which should be modified, and those which could be discontinued. An elaborate system of subcommittees was created to review various types of questionnaires. Industry representatives were invited to visit Washington and, in collaboration with the Committee, to make a review of each repetitive reporting form. Further, the Committee solicited criticism and suggestions from over 1,000 trade associations and their members in a drive to ferret out unnecessary duplication, to define areas where consolidation of reports could be effected, and to restudy the reports which caused difficult filing problems. In response to this request, a total of approximately 40,000 letters were received in mid-1942 embodying the important complaints, criticisms, and suggestions of industry.

The vast energy expended and the calibre of talent used produced disappointing results. These were made the poorer by exaggerated claims of success by the Committee. The Committee itself was responsible for relatively few revisions or discontinuances. Most revisions and discontinuances claimed had been planned by the WPB divisions before the Committee was established. Pressures generated by industry criticism and the existence of the Committee no doubt hurried the process but the fact remains that relatively few revisions or discontinuances resulted from the literally blinding light focused on each inquiry. These facts support the conclusion that the complex and burdensome reporting system was the result not of poor questionnaire control, inadequate policing, or the absence of self discipline, but rather of confused basic WPB policies. And despite the vast reporting structure, there is reason to believe that the WPB erred on the side of too little information rather than on too much.

The quality of industry's critical response was also disappointing. Analysis of the thousands of letters from industry demonstrated that the burden of the WPB requests on industry, and industry objection to reporting forms, centered almost exclusively on the basic material report forms used under the individual allocation systems. The principal difficulty of industry in filing these forms resulted from the request for end use information which industry neither had nor could readily get. Of the hundreds of forms in use, only a handful outside of the raw material control area were the subject of industry comment. The trouble lay not in the report forms themselves but in the material control systems which they implemented. Most of the report forms in use were needed to facilitate operations under the various overlapping and inconsistent control systems in force. Before order could be brought into the reporting systém there had to be a definite policy established for control of material flows. The coordinated efforts of industry in criticizing the WPB questionnaires added practically nothing to what was already known to statistical technicians and the top staff of the WPB.

Even though the Committee was less successful than its claims, it performed a valuable function. It focused the attention of businessmen both in the WPB and in industry on the problems of wartime informational requirements, especially the character of the information needed by government and the ability of industry to supply the data. As a result a decision was made to establish on a permanent basis the machinery through which the viewpoints of industry could be represented in the issuance and continuance of the WPB's reporting forms.

In October 1942 the permanent position of Chairman of the Committee on Data Requests from Industry was established and filled by a business man. This function continued well into 1944 when, as a result of a sharp decline in the introduction of new data requests, the position was abolished.

ADVISORY COMMITTEE ON GOVERNMENT QUESTIONNAIRES

In addition, by a letter dated October 9, 1942, the Director of the Bureau of the Budget requested the President of the Controller's Institute of America to establish an Advisory Committee on Government Questionnaires composed of five members, one representing each of the following industry organizations: the National Association of Manufacturers, the American Trade Association Executives, The Controllers' Institute of America, the Chamber of Commerce of the United States, and the American Retail Federation. The Committee was asked to take steps necessary "to make the advisory services of business available to Government agencies . . . in promoting the improvement, development and coordination of all federal statistical services." The Committee, later expanded to include representation from the National Association of Commercial Organization Secretaries, and the National Industrial Council, was independent of government and responsible only to the business community. Its functions were purely advisory.

The Committee operated in large part through subcommittees. Some groups restricted their work exclusively to a single industry, as for example questionnaire problems of the arc welding industry. In other cases, subcommittees represented production segments of an industry, such as the organic chemical industry, or a functional part of an industry, as for example problems of accounting in questionnaires of the public utility industry. In still other cases subcommittees devoted their attention to general problems that affected many lines of business, as for example industry's problems in filing WPB-732. Subcommittees met regularly and devoted themselves to the intimate details of reporting problems. Panels were created in emergencies to consider nonrecurring problems. Exclusive of such panels, 58 subcommittees were appointed during the first year of the Advisory Committee's existence.

The Advisory Committee performed a valuable service to government, a service not measurable in quantitative terms. During its first year of operation some 300 specific recommendations were made to the Bureau of the Budget and substantially all were accepted both by the Bureau and the agencies whose programs were affected. There were also hundreds of unrecorded instances where questionnaires were voluntarily revised and substantially improved by persons and agencies affected before their original issuance. In addition, as noted by the Director of the Bureau of the Budget in a letter to the Chairman of the Advisory Committee on Government Questionnaires, dated December 3, 1943, its existence encouraged greater care in the development of governmental forms and greater attention to the abilities of respondents to supply the information requested. Although these considerations applied to all Government agencies, the bulk of the work of the Committee pertained to questionnaires of the WPB since its surveys loomed so large in the total of all Government inquiries.

TASK COMMITTEE ON REDUCTION OF SPECIAL APPLICATIONS

In mid-1943 renewed industry complaints about the WPB's informational demands coincided with a general over-all improvement in supply-demand relationships in many basic materials and intermediary products. These circumstances made it difficult to visualize the need for continuing many reports. A Task Committee on Reduction of Special Applications was therefore established by the Operations Vice Chairman and the Program Vice Chairman to determine the minimum number of special applications that should be required from industry. Primary responsibility for a general housecleaning was placed upon the WPB Division Directors. Since the Chairman of the Task Committee was also the Deputy Vice Chairman for Operations with full authority to act when a decision was made, prompt results were expected.

A reduction of 40 per cent in the applications required from industry was set as the objective to be achieved through the following methods:

"1. Elimination of allocation or special authorization procedures wherever supply warrants.

2. Substitution of limitation orders or restrictive provisions in orders for those which now require special applications.

3. Small order exemptions to be incorporated in the existing orders so that not less than 10 per cent of supply will be distributed in this manner unless conclusive evidence is given justifying a lower percentage.

4. Replacement of consumer applications by suppliers' order board reports.

5. Substitution of quarterly applications for monthly applications.

6. Substitution of quotas for special applications (including MRO) now required wherever consumption and use have been established."

Coincident with work on weeding out unnecessary application forms, the Task Committee was to attempt standardization wherever possible, particularly as to end-use information; to ask only for facts readily available to manufacturers and eliminate those questions which manufacturers had to answer from supplier or customer records; and to reduce wherever possible paper work other than questionnaires.

COMMITTEE ON DEMOBILIZATION OF CONTROLS AFTER V-J DAY

In the summer of 1944, when the collapse of German resistance seemed to be imminent, the pace of reconversion preparations, which had been in progress from as far back as the spring of 1943, accelerated quickly. Leading the reconversion plans was a Committee on Demobilization of Controls after V-E Day, familarly known as CODCAVE. When the Battle of the Bulge in the winter of 1944–45 shifted attention from reconversion to renewed war-production, a new Committee on Period One (CPO) was established in February 1945 to review the program developed by CODCAVE and keep it up to date. This Committee functioned until V-J Day. When Germany collapsed it had a complete program for relaxation of controls ready to be put into effect at once. Despite the unexpectedly short period between V-E Day and V-J Day, it had a similarly complete plan when the Japanese surrendered.

The fundamental concern of these Committees, of course, related to basic production and distribution controls. But accompanying those controls were reporting instruments. Plans concerning questionnaires were developed by a Sub-Committee on Reports and Data Requests. As a part of the work of the principal Committee, every control policy, order and regulation was reviewed by the subcommittee.

The principal objectives which the subcommittee established in reviewing questionnaires are worthy of note. First, of course, was the desire to eliminate all non-essential reports growing out of relaxation of basic controls and management needs. Second, it was decided to maintain an intelligent reporting program which would permit the WPB and other war and reconversion agencies to be properly informed and prepared to take any necessary action in the reconversion period. Third, it was considered important to preserve in an uninterrupted series all reports which could be regarded as having post-war value to permanent Government agencies. In contrast with the approach to comprehensive elimination of restrictive controls, it was felt that in revising its reporting requirements following V-E Day, the WPB should err on the side of caution rather than over-hasty relaxation. The experience of five years of struggle in building up a reporting system and in fully understanding the strategic role which it played in management was not lost. It was considered to be good policy for the WPB to keep its "steering gear and brakes in good condition," a function which could be performed by the maintenance of minimum reporting.

Preparations made on the basis of these principles resulted in a substantial reduction of questionnaires upon the arrival of V-E Day and a further substantial reduction on V-J Day. The reductions were orderly and resulted in the maintenance of reporting instruments which proved to be of great value in the post-war reconversion period.

CENTRAL REVIEW PROBLEMS AND POLICIES

The core of the central review problem rested in the effort to satisfy all competing pressures present at the point of central review and at the same time gradually raise technical levels of statistical proficiency and sound operating procedures. Each data request originally had to be reviewed not only as a technical and professional problem but also in the light of a wide variety of complex influences growing out of the nature of the WPB organization, the personalities of administrators, the status of various control mechanisms, the degree of statistical coordination present, the urgency of data requirements, present and prospective control policy, industry's ability and willingness to comply, the nature of other comparable questionnaires and literally dozens of other such political, administrative, and technical issues. And once a decision was made it could not be forgotten. Review of past actions was a necessary accompaniment to original control, for the WPB divisions often maintained and supported questionnaires whose usefulness had ended or whose efficiency could be improved by the application of newly tested technical methods.

Fundamental in the policy of form clearance was the idea that better results would be obtained in the long run if emphasis were laid not so much upon rigid adherence to fixed procedures as upon the personalities concerned with questionnaire formulation and data use. WPB's personnel was a heterogeneous group of administrators, with varying degrees of background in methods of data collection and use, and with substantial authority and responsibility in the management of individual products and materials. It was natural for these individuals to insist that, upon the basis of their knowledge of their own industry and war needs, a survey developed by them not only was necessary but also administratively practical. It was hard for central clearance to argue against the feeling that "it is necessary to have this information in this way to fulfill the responsibilities of this office." Reconciliation of such ideas with technical standards and industry's ability to comply was often not easy.

The approach to such questions was conciliatory. The goal was to nurture acceptance of the idea that questionnaire clearance existed to aid administrators to solve reporting problems, to harmonize their methods with others in the WPB, and to utilize experience of others in making the solution of their problems easier. It was a policy of aid rather than dictation.

Such policies were effective. They generally engendered a feeling of reason into discussions of problems which otherwise might have been settled arbitrarily with resulting appeals to higher authority and bitter conflict.

Neither administrators nor industry advisory groups were always thoroughly familiar with the record-keeping practices of industry from which data requests had to be filled. There were many instances of violent critical reaction from industry to forms which had been approved by the WPB and industry advisory committees. In some instances forms had to be withdrawn because industry was in no position to respond. In others, special "Task Groups" had to be established to simplify questionnaires before the great majority of respondents could be expected to comply. There were many instances in which those who should have had knowledge of their industry's ability to file reported information were responsible for data requests which could not be filed. Often, against the limited knowledge of sponsors of questionnaires, clearance had to array information concerning the ability of industry to comply.

In addition, of course, were the problems of matching the proposed questionnaire against technical standards, acceptable statistical methodology, the WPB's control policies, and sound administrative procedure. From the beginning it was accepted practice in the clearance function to assume that those responsible for policy and administrative action should themselves define their information needs. In view of this policy, form clearance devoted itself to solving those problems which might jeopardize the collection, tabulation, and use of required statistics. But this was not always easy and the line between conciliation and dictation often was blurred.

Important in the functioning of centralized form clearance was the adoption of technical statistical standards and methods to current questionnaire problems. The form clearance problem was one of watching experience throughout the WPB and applying that experience to the acquisition of information needed for management. The history of form clearance in this regard was an uphill climb in instituting accepted methods in connection with the use of "cut-off" points, standardization of forms applicable to general orders, gearing data requests to industry record-keeping practices, eliminating "interesting figures" not essential for the purposes of the questionnaire, and generating a freer flow of information from one division to another to avoid duplication and overlapping of data requests. Many desirable technical methods could not be introduced arbitrarily but through a process of education their incorporation resulted in a gradual improvement in questionnaires. As experience brought with it assurance of rightfulness there were many technical standards applicable to surveys and questionnaires that had to be met before form issuance could be approved.

Relating reporting to policy raised a dual problem. On the one hand, when a policy was established which resulted in questionnaires with which industry could not comply an effort had to be made to alter the policy. On the other hand, when a policy was established and a questionnaire contradicted that policy the questionnaire had to be modified or rejected. Implications of these problems were far-reaching.

Policies connected with early individual allocation systems illustrate the first type of problem. As discussed previously, most of the early allocation systems relied heavily upon acquisition of information which it was often exceedingly difficult and sometimes impossible for industry to report. The problem was attacked in a number of ways. First, an effort was made to develop in the WPB an understanding of the difficulties in industry in reporting end-use information. Increasing understanding eventually led to requests for types of end-use information generally available in the records of respondents. As a result, many so-called end-use systems developed into use systems. In many areas, however, as noted elsewhere, such results were not achieved. Second, an effort was made to prevent the rise of numerous new allotment systems which conflicted with those in use. Although form clearance did not have authority to correct or initiate fundamental policy it did have authority to disapprove new questionnaires. Judicious exercise of this power resulted in discouraging many new allotment schemes. Third, strenuous efforts were made to eliminate duplication and overlapping of reporting by the modification of existing questionnaires supporting the systems. The eventual effect was to crystallize control policy and coordinate many individual allocation systems into workable patterns.

Tying reporting to established policy was often more difficult when basic policy was itself confused and overlapping. In such circumstances, relating questionnaires to policy was not a simple undertaking. The coterminous existence in 1942 of the Production Requirements Plan and the many individual allocation systems is a case in point. Considerable conflict and duplication existed among these systems and provided a fertile field for form clearance to launch a unification program. Effective corrective measures, however, were impossible in the absence of clear-cut top policy. Means for taking some corrective measures came only with clarification of basic metal allocation policy upon the announcement of the Controlled Materials Plan.

Developments relating to scheduling of components provide another illustration of the way in which reporting standards were sometimes subordinated and thwarted by ill-defined basic policy. Before the adoption of M-293 (the General Scheduling Order) various WPB divisions had instituted action to request manufacturers of fabricated components to report their individual orders. The purpose of such questionnaires was to place the WPB in a position to direct shipment of a specific item to a specific customer. Such a procedure not only was contrary to the basic policy of the priority system as it then existed but necessitated more personnel to handle the mass of data which would be collected than the divisions possessed. In addition, any benefits which might have been obtained were small when compared with the tremendous reporting burden such questionnaires would have created for industry. With few exceptions, these requests were therefore disapproved.

Issuance of M-293 early in 1943 altered WPB basic policy and also clearance policy with respect to scheduling forms. Central clearance recognized the fact that the mass of information required by M-293 could hardly be used in a fashion to justify the heavy reporting burden imposed on industry. But central clearance, in the face of firm WPB policy, was helpless to disapprove the questionnaires for such reasons.

Estimates of the impact of WPB questionnaires on industry were always considered in the review of questionnaires. The establishment of committees to study the problem has already been discussed. Beyond the problem of eliminating needless questionnaires lay the more intricate problem of adapting questionnaires themselves to fit industry records. Interest in this problem was not entirely the result of concern for industry. By facilitating industrial response the WPB could obtain more accurate, complete, and prompt response to data requests.

Attention had to be given to programs of other war agencies and the utilization of other agencies in both the mailing and tabulating of questionnaires. The clearance group sought to coordinate, directly and through the Bureau of the Budget, the WPB reporting program with that of other Federal agencies. In addition, arrangements were made with the Bureau of Mines, the Bureau of Labor Statistics, the Bureau of the Census, and other agencies to modify their regular reports to meet special needs of the WPB. In some cases, WPB orders made mandatory the filing of another agency's report since that was the easiest method to obtain necessary information.

Of much importance in the WPB's operations was the use of other agencies in collecting and tabulating WPB reports. These agencies made available to the WPB technical skills in coding, editing, and tabulation questionnaires. Arrangements for acquiring these services were made through the clearance group.

BOOK REVIEWS

Edited by
Oscar Krisen Buros
Rutgers University

Experimental Designs in Sociological Research. *F. Stuart Chapin* (Professor of Sociology, Chairman of the Department, and Director of the School of Social Work, University of Minnesota, Minneapolis, Minn.). New York 16: Harper & Brothers (49 East 33rd St.). 1947. Pp. xi, 206. $3.00.

Review by O. Kempthorne
Associate Professor of Statistics, Iowa State College
Ames, Iowa

In the introduction the author states "the newly developed research methods of experimental design, applied in the community situation, seem to point the way by which we can overcome, in the course of time, some of the complexities of social interaction that have hitherto baffled rational and objective description of certain acute problems of human relations." There can be no disagreement with this statement; and, in fact, one could well take the viewpoint that until the general methods of experimental design as applied, for example, to biological research are part of the approach to the investigation of social situations very little progress can be made. It is unfortunate then that the author has gathered together in this book, as noteworthy examples of the use of experimental designs in sociological research, nine studies, only one of which would satisfy the elementary basic requirements of a modern experiment. This statement will be dealt with later in the review. First, however, it would be well for the reviewer to state that he realizes the difficulties which arise in any attempt to apply to sociological research the principles of experimental design as laid down for example by R. A. Fisher in *The Design of Experiments*. It is very easy for the agronomist to divide his experimental area into plots to which treatments are applied at random, whereas the division of a community into several parts for random application of sociological treatments, such as varying types of housing, is extremely difficult. It is possible to visualize the application of sociological treatments to communities selected at random, though the pressure of public opinion may prevent research workers from adopting this procedure for any important problems. Professor Chapin discusses these difficulties at some length throughout the book, and the biologist or agronomist can only feel happy that his problems are so much simpler.

The chapters of the book are entitled: (1) Natural Social Experiments by Trial and Error; (2) Three Experimental Designs for Controlled Observation; (3) Cross-Sectional Experimental Design; (4) Projected Experimental Design: The Classical Pattern of "Before" and "After" Experiments That Operate from the Present to the Future; (5) Ex Post Facto Experimental Design: From Present to Past; (6) Sociometric Scales Available for Control and the Measurement of Effects; and (7) Some Fundamental Problems and Limitations to Study by Experimental Designs. This is an imposing array of chapter headings with four different types of experiments mentioned, natural, cross-sectional, projected, and ex post facto. The question arises immediately of what the author understands by the words "experiment" and "experimental design." The word "experiment" has been used presumably for centuries to describe the taking of observations according to some plan, though in the biological sciences it now connotes the application and comparison of various different treatments or technical procedures. The necessity for clear thinking on this point has been emphasized in a recent paper* where the author differentiates between two types of experiment, the one for example to measure an absolute constant such as the velocity of light and the other for example to evaluate the effect of a special dietary supplement. The usage of the word "experimental design" is well established by now to mean a plan for performing a comparative experiment. This implies that various treatments are actually applied by the investigator and are not just treatments that happened to have been applied to particular units for some reason, known or unknown, before the "experiment" was planned. This condition rules out practically all of the experiments and experimental designs discussed by the author. This so-called "natural" experiment should be described as a phenomenon and has little inductive value; for it is well established that the same set of environmental conditions will not produce the same effect in all cases. The "cross-sectional" designs discussed have a superficial appearance of validity, and would be completely valid if (and only if) a randomization procedure had been adopted in assigning treatments. The fact that pairs of individuals have been matched on several control factors does not mean that they are identical: matching of experimental units is the normal procedure in all types of experiment, but is useless without randomization and replication by which the closeness of the matching may be estimated. Experimental errors are frequently quoted in the book when there is neither randomization nor replication of the treatments. The author appears (Chap. 7 and elsewhere) to be aware of these considerations, which if given their proper weight would have resulted in the exclusion of practically all the studies on the grounds that they are examples of the misuse of basic concepts and are incapable of unbiassed interpretation. The chapter on "Projected Experimental Design" contains the only one of the nine studies in which the principle of randomization was used, namely Dr. Reuben Hill's study on "Staff Stimulation to Social Participation and Social Adjustment." This study was not, however, entirely satisfactory because of the population of 306 freshmen considered, there were only 171 for whom complete records were available. It appears that what the author terms a projected design is satisfactory if the concepts of randomization and replication are used properly. The reviewer fails to understand how the "Ex Post Facto Experimental Design" can be given a logical justification. An example of this "design" was the determination of the relationship of public housing in New Haven to juvenile delinquency, in which the earlier records of 317 families who had lived for a few years in one of the Developments of the New Haven Housing Authority were compared to their records during these few years. Some of the difficulties in the interpretation of such data are obvious to anyone experienced in the analysis of data and there is no need to discuss them.

---

* Anscombe, F. J. "On the Validity of Comparative Experiments" *J Royal Stat Soc* (In press).

There appears to be some confusion in the statistical literature on the differences between experimental designs and survey designs. Both involve patterns by which observations are taken, but in experimental designs the objective is to estimate the effects of applied treatments and in survey work the objective is to estimate certain characteristics for a particular defined population. There are possibilities of using experimental design patterns in survey work, as for example the use of deep stratification with a Latin Square. Again, if the purpose of the survey is to establish the existence of associations in the population, a sample that is balanced on some factors may be useful. Causal relationships are frequently inferred from such associations, but they cannot be regarded as having the sound scientific basis of inferences drawn from experimental data.

Apart from the above objections which are in themselves sufficient to make the book unreliable if not dangerous for workers in social sciences, there are very many instances of imprecise thinking and writing. Two examples are sufficient evidence: "The conventional tests of the significance of sampling are based upon the theory of random samples, and in the present stage of experimental work, as we have attempted to show, it is the terminal homogeneity and purity in the sample, rather than initial representativeness of heterogeneity, that is important in demonstrating the *real* relationship between the two variable factors" (p. 104), and "Furthermore, since the critical ratio of 2.90 is computed from a difference between extremely pure samples, its significance is much enhanced by this fact" (p. 107). (The "critical ratio" appears to be Student's "t".)

The author in spite of numerous references and discussions of the question does not appear to understand the concepts of prediction and causation. Mathematical economists have for some time differentiated between the two problems of predicting events and estimating structural relationships. The interpretation of correlation coefficients appears to the reviewer to be particularly vague.

Further general criticisms of the book are that the author takes as the fundamental rule of the experimental method the varying of only one condition at a time with all other factors maintained constant (p. 1), and there is lack of clear discussion of the formulation of hypotheses in an unambiguous way before looking at the experimental data, this formulation of course including the definition of some population for which the conclusions are to be relevant.

Principles of Medical Statistics, Fourth Edition. *A. Bradford Hill* (Professor of Medical Statistics and Director of the Department, London School of Hygiene and Tropical Medicine, University of London, London, England). London W.C.2: Lancet Ltd. (7 Adams St., Adelphia), 1948. Pp. xi, 252. 10*s*. 6*d*.

REVIEW BY J. YERUSHALMY
Professor of Biostatistics, University of California
Berkeley, California

MANY workers in clinical and preventive medicine have found the previous editions of this book the most satisfactory source for acquiring a working knowledge in statistics and an understanding of the role of quantitative methods in medicine and public health. The publication of the fourth edition is evidence at once that the demand for knowledge in this field is present in an increasing number of members of the medical professions and that this book is singularly effective in satisfying this demand.

Of greatest interest to professional statisticians are the techniques of presentation which accomplish the difficult task of making statistical methods palatable to a nonmathematical audience. These techniques, in the opinion of this reviewer, derive their effectiveness from a sound evaluation by the author of the intellectual level of the readers for whom the book is intended. It is apparent that Professor Hill addresses himself to an audience of intelligent and capable persons in their own fields of endeavor, even if they are lacking in mathematical training. Consequently he does not "talk down" to them; rather he introduces the statistical concepts in the framework of medicine and public health with which his readers are familiar. He thus succeeds in developing the statistical techniques without the use of mathematical symbols and formulae in such a way that the rationale behind them is comprehended. The concepts become acceptable as logical entities although no mathematical derivations are presented.

One minor exception to the generally lucid and comprehensive method of presentation may be mentioned. In discussing the standard error of the difference between two sample proportions Professor Hill presents the usual two formulae, one in which an estimate of the universe proportion is obtained from a combination of the two samples, the other in which individual proportions of the two samples are used in the determination of the standard error of the difference. The explanation of the later technique, however, is presented in terms of two different estimates of the standard error of the proportion in the universe (pp. 111–112). It would probably be more easily acceptable to the reader if the hypothesis of a single universe was not retained in the second approach.

In this fourth edition Professor Hill has substantially elaborated and strengthened many of the weaker points of the previous editions. Tabular and graphic presentation and the description and calculations of averages is more adequately covered. Sampling is discussed more adequately and the concept of normality is presented more clearly. The development of the life table and its functions is particularly lucid. The chapters on sampling, averages, proportions, differences, and chi square are enriched by numerous examples, many of them in the field of public health. There are also additions in the number of the very instructive examples of fallacies, difficulties, and problems of selection.

In spite of the additions and elaborations of so many topics, Professor Hill succeeded nevertheless in his objective to have the book "kept within limited bounds and confined to 'arithmetic guided by logic'."

The Teaching of Statistics: A Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics. *Harold Hotelling*, Chairman (Professor of Mathematical Statistics and Head of the Department; Associate Director, Institute of Statistics; University of North Carolina), *Walter Bartky* (Dean of the Division of the Physical Sciences and Professor of Applied Mathematics, University of Chicago, Chicago, Ill.), *W. Edwards Deming* (Adviser in Sampling, Bureau of the Budget, Washington, D. C.), *Milton Friedman* (Associate Professor of Economics, University of Chicago, Chicago, Ill.), and *Paul Hoel* (Associate Professor of Mathematics, University of California, Los Angeles, Calif.). Reprinted from *Ann Math Stat* 19(1):95–115 Je '47. Washington 25, D. C.: National Research Council (2101 Constitution Ave.), 1948. Pp. 95–115. Paper.

The Teaching of Statistics in Universities and Colleges. Based upon a report by the Royal Statistical Society's Committee on the teaching of statistics: *E. S. Pearson* (Chairman), *R. G. D. Allen*, *H. Campion*, *William Elderton*, *C. Oswald George*, *R. F. George*, *M. Greenwood*, *D. Heron*, *J. O. Irwin*, *H. Leak*, *E. C. Rhodes*, *R. Stone*, *L. H. C. Tippett*, and *J. Wishart*. Preprint of *J Royal Stat Soc* 110(1):51–7 pt 1 '47. London W.C.2: Royal Statistical Society (4 Portugal St.), 1947. Pp. 7. Paper. 6*d.*

REVIEWED BY TRUMAN L. KELLEY
Professor of Education, Harvard University
Cambridge, Massachusetts

AS THESE reports cover the same ground, it is advantageous to compare their points of view and recommendations. The British report is the more realistic in that it starts with the status quo in the United Kingdom and Eire, where the teaching of statistics is not nearly as extensive as in the United States, and makes recommendations which may, perhaps, be attainable under the conflicting academic pressures. The American Report is idealistic. It states with conviction and well-documented argument what ought to be the situation, but it pays scant attention to the sundry historical origins of statistical development and teaching, or to the multitudinous reasons for specialization and differentiation.

The British report states that "There are few more urgent needs today than the increase of the supply of first class statisticians," and the American report repeatedly expresses the same idea.

The British committee asserts that social science workers require a knowledge of "descriptive statistics." There is no specific reference to this initial and elementary need in the American report. This matter is of importance, for the American committee believes that a knowledge of basic statistical principles is a part of a liberal education which all college students should get. This basic need is acquaintance "with the broad principles of inductive inference." It seems to the reviewer that there is implicit in the British recommendation an induction of the student into statistics via the subject matter of his field of specialization, and in the American an induction via logic, including principles of mathematics and probability. It is needless to say that these approaches are far asunder. By the one approach the student is initially aware of virginal data and with later study acquires the power of correctly appraising it, while the other approach would see that the student first acquires the powers and the tools for sound appraisal and then applies them to the field of his choice.

This second approach seems necessary if all students, irrespective of their variety of interests, become acquainted with statistics through a general and comprehensive orientation course. Both committees favor the inclusion of statistics (of a sort not clearly defined by the British committee) as a minor part of a universal basic course.

Both committees believe, as stated by the British committee, that: the statistical approach is so fundamental to the modern way of looking at things . . . that it should form a part of the mental equipment of the educated man . . . " This committee adds, with refreshing realism and modesty, that " . . . we do not propose at present that the universities should attempt to fill this gap in general education . . . " However, both committees attack the problem, for they propose statistical courses, and the American committee goes into detail in recommending the type of instructor that is needed.

At the lowest level the British committee advocates a general course to familiarize students with statistical ideas and the most commonly used elementary methods, and the American committee asserts that "the fundamental logic and philosophy of statistics can be taught at an early stage." This is to include the broad principles of inductive inference, concepts of sampling variation, randomness, and statistical predictability, the difference between inductive and deductive statements, the nature of statistical estimation, and the nature of statistical hypothesis. As a teacher who has tried a variety of approaches, the reviewer recommends one which is not clearly within the scope of either of these committee recommendations—it is via data already within the field of interest and acquaintance of the student. The phenomena of social and scientific experience constitute the starting point. Their need for interpretation through sound logic is so obvious and so personally felt that there is no sense of unreality and no lack in motivation. Because of the variety of interest, a single introductory course does not suffice so, contingent upon sufficient registrants, there may well be introductory courses in statistics in biology, agriculture, economics, education and psychology, physical sciences, etc. Such an approach is vigorously castigated by the American committee, which advocates an introductory course under the aegis of the department of mathematics (expanded, to be sure, to include an expert statistician) and which even goes so far as to imply that the present situation wherein many departmental introductory courses are given by semi-trained statisticians is worse than if no courses at all were offered.

Both committees admit the paucity of adequately trained teacher personnel. In the light of it we must, as a practical matter, consider the benefits derivable from a course given, say, in the college department of economics by (*a*) a well-trained mathematician and statistician who is quite ignorant of economic issues, (*b*) a well-trained economist who is quite ignorant of statistical niceties, or (*c*) one semitrained in both statistics and in economics. The reviewer observes that situations change and that with national standards in the various disciplines set by such men as Fisher, Hotelling, Pearson, etc. a second giving of a course by an (*a*), a (*b*), or a (*c*) will migrate from a first giving in a commendable direction. We can say that two decades of American procedure that has been contrary to that endorsed by the American committee has resulted in a truly remarkable advance in experimental and applied statistics. The British committee implies what the decades have proven, that high competence in statistics has not been dependent upon college courses in statistics. Such courses certainly should broaden the base from which individual genius can rise to an apex, but let us be wary of a definition of path which attempts to fix the initial steps or the terminal product.

The reviewer gathers that the American committee pictures statistical training through the following sequence: (1) a first course, or part of a course, for all students, given by a mathematical statistician; (2) a course for "future consumers" given by an applied statistician; (3) a course for "future users of statistical methods" which is in part mathematical statistics and in part applied statistics; and (4) a course for "future research workers and teachers . . . ," given under the original parental roof of mathematical statistics. The British recommendations are not so specific, but they write, whether with approval or regret is not clear, that "it seems inevitable that there should be separate courses for . . . students specializing in various sciences." They also say that "the necessity, at this level, of supplementing lectures by practical work cannot be too strongly emphasized." The necessity for a number of centers for advanced courses is emphasized. The sort of advanced course here implied seems quite similar to number (4) of the American committee scheme, for they write "the basis of statistical theory and its ally the theory of probability is essentially mathematical."

The American committee, by omission and by inclusion, reveals what it considers to be preparatory background for students of statistics. It at no point cites knowledge of data in some scientific field as essential. It urges knowledge of the calculus—making the happy suggestion that it is appropriate for high school study—but fails to emphasize matrix and other advanced algebra which are much more important for statistics. In the judgment of the reviewer, its greatest failure is in not building upon the usual normal psychological course of student development—(*a*) knowledge of general affairs, including a smattering of mathematics, (*b*) a developing interest in some scientific field, (*c*) an epochal experience when data are discovered to be an original source of new knowledge, (*d*) a determination to utilize this source and to acquaint oneself with the techniques for making it reveal its hidden information. Not until this last is tapped can we expect that men will devote themselves to the training necessary to become great statisticians and experimentalists. The reviewer believes that the end product in the teaching of statistics should transcend, if not in height certainly in breadth, that implied in the term "mathematical statistician."

The American committee deplores the general lack of mathematical competence of most teachers of statistics in different subject matter fields. This is deplorable as is their lack of knowledge of the genius of data in their fields. However, the progress of recent decades should make one optimistic, and these two committee reports should encourage college presidents to strengthen and broaden the instruction in both mathematical and applied statistics.

Nomography. *A. S. Levens* (Associate Professor of Mechanical Engineering, University of California, Berkeley, Calif.) New York 16: John Wiley & Sons, Inc. (440 Fourth Ave.), 1948. Pp. viii, 176. $3.00.

REVIEW BY M. KAC
Associate Professor of Mathematics, Cornell University
Ithaca, New York

THIS book is devoted to an elementary presentation of the basic principles used in design of nomograms (alignment charts). The material is standard but attractively presented and illustrated by numerous examples.

The mathematical prerequisites are modest and consist of no more than elementary algebra (through logarithms) and plane geometry.

The book is an outgrowth of nomography courses for engineering students and it is therefore only natural to find most of the applications and illustrations drawn from that field. There is no doubt that nomograms are useful as a computational aid in many engineering problems. The reviewer doubts however, that nomography will prove very useful to statisticians. In fact, out of thirty-odd nomograms collected in the Appendix and which in the words of the author "may prove useful in the fields of engineering, production, business, and statistics" there are only two taken from the field of statistics. These are a nomogram for the standard deviation (p. 170).

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n}}$$

and a nomogram for the correlation coefficient (p. 171)

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\ \sqrt{\Sigma y^2}}$$

Both of these nomograms can be used only after Σx^2, Σy^2, Σxy have been calculated. Since the computation of Σx^2, Σy^2, Σxy is by far the most laborious part, it seems a little naive to expect that the above nomograms will introduce a significant saving of time.

Foundations of Economic Analysis. *Paul Anthony Samuelson* (Professor of Economics, Massachusetts Institute of Technology, Cambridge, Mass.). Harvard Economic Studies, Vol. 80. Cambridge 38, Mass.: Harvard University Press (38 Quincy Street), 1947. Pp. xii, 447. $7.50.

REVIEW BY GERHARD TINTNER
Research Associate, Department of Applied Economics
Cambridge University, Cambridge, England
ON LEAVE: *Professor of Economics and Mathematics*
Iowa State College, Ames, Iowa

THIS is a very important contribution in the field of pure economics which ought to be of interest to econometricians and those statisticians who concern themselves with economic matters.

The author endeavors throughout to derive "operationally meaningful" propositions in economics from the assumption of maximizing behavior. Operationally meaningful statements are hypotheses about empirical data which could conceivably be refuted, if only under ideal conditions. These ideas are related to those of some modern positivist philosophers, especially Bridgeman. It is perhaps unfortunate that the author has not included a methodological chapter in his book. An explicit statement of his philosophical ideas may have contributed to the understanding of his book. It may have helped some economists who are sceptical about the applicability of radical positivism and behaviorism in the social sciences.

The first part of the book is concerned with economic statics. It also includes a chapter on welfare economics, which seems a little out of place. The theory of maxima in all its ramifications is treated at great length. There is also a more systematic mathematical appendix on this subject.

On the basis of a very general approach the theory of cost and production and also the theory of consumer behavior is ably stated. An interesting feature in the treatment is the discussion of maxima if the functions in question are not continuous. This is a very important advance, though the discontinuities considered may not be the most important ones met with in economic life.

Monopoly and monopolistic competition get very little attention, to say nothing of more complicated situations like duopoly and bilateral monopoly. This narrowness of the point of view makes the book a little less useful than it might have been.

A short section on the economic theory of index numbers (pp. 146–156) should be of particular interest to the statistician. It is probably the best short statement available about the modern theory of cost of living index numbers.

A very interesting chapter on welfare economics concludes this first section of the book. In its brief compass it contains most of the important new results of the modern theory in this field.

The second part of the book deals with economic dynamics. But the discussion is almost entirely based upon purely formal consideration of equilibria and their relation to comparative statics. The "correspondence" principle of the author states that "the problem of stability of equilibrium is intimately tied up with the problem of deriving fruitful theorems in comparative statics" (p. 258). The author succeeds in reaching very interesting conclusions in this way. His work is inspired by the modern theory of dynamics (in physics), especially by the work of George D. Birkhoff (*Dynamical Systems*, 1927). A mathematical appendix deals with the theory of difference and other functional equations. The examples given in this section of the book include also a very interesting "dynamization" of the Keynesian system. But it appears here and elsewhere that not many definite conclusions can be drawn, even after making a number of somewhat arbitrary assumptions.

The second part of the book is hence somewhat disappointing to the economist. It is evident that not many valid and interesting conclusions can be drawn from such narrow assumptions as made by the author. The whole theory of anticipations and expectations is either neglected or appears only incidentally. Uncertainty hardly gets an adequate treatment in the small space allotted to it. Samuelson's book is in this respect inferior to the earlier work of J. R. Hicks (*Value and Capital*, 1939) and others who have endeavored to deal with problems of economic risk and uncertainty. I believe that especially the theory of formation of anticipations forms a true link between static economics and the more useful and interesting economic dynamics.

Chapter 11 deals with the classification of equilibria. This discussion is again based upon modern (physical) dynamics. The usefulness of the distinctions made in economics is not immediately apparent. The subsequent discussion of the business cycle is very short and not on the high level maintained in the earlier parts of the book.

The author devotes towards the end of the book a few pages to the theory of stochastic systems. The treatment is much too brief to yield important results. Some of the ideas presented may however prove stimulating to the statisticians working in this field, especially the indications about nonlinear stochastic systems.

In summary, the book ought to be recommended very highly to the mature reader who possesses an adequate knowledge of modern economic theory and higher mathematics. A conclusion which he may draw from its contents is the following: Assuming only maximizing behavior (something like the "economic principle" of the elementary text books) very few interesting conclusions about economic behavior can be drawn even by so brilliant a theoretician as Samuelson. There is a great deal of valid and important economic theory contained in his book. The reader might be inspired by the comparative failure of pure economics to devote his interest to a more promising approach. Econometric investigations combine modern economic theory, which is so ably presented in this book, with modern statistical methods in order to derive valid empirical conclusions.

STATISTICAL METHODOLOGY INDEX, NO. 13

A QUARTERLY GUIDE TO CURRENT LITERATURE

Edited by
OSCAR KRISEN BUROS
Rutgers University

This bibliographic service presents statistical methodology literature (articles, books, theses, and chapters) published in 1945 to date. Anonymous references are listed first; other references follow in alphabetical sequence by first-named authors. Volume number, issue number (in parentheses), pagination, and date of issue are given for each journal reference. Stars indicate books, theses, and pamphlets; asterisks indicate publications seen by the editor of this INDEX. *Authors of papers on statistical methodology are invited to send reprints to Oscar K. Buros, Rutgers University, New Brunswick, New Jersey in order to facilitate prompt and accurate listings. Foreign-language papers should be accompanied by English translations of titles. Information is desired about references omitted.*

1952. **Calculating Machines.** *Nature* 161(4097):712–3 My 8 '48.*

1953. **Statistics and Research.** Editorial. *Chem Age* 58(1488):96 Ja 17 '48.*

1954. ADLER, H. A. (Senior Engineer, Engineering Department, Line Design Division, Adams St., Chicago 90, Ill.), and MILLER, K. W. (Assistant Director, Armour Research Foundation, Illinois Institute of Technology, Chicago 16, Ill.). **A New Approach to Probability Problems in Electrical Engineering.** *Trans Am Inst Elec Eng* 65:630–2, discussion 1118–9 '46.* Same as 1881 plus discussion.

1955. ARROW, KENNETH J. (Research Associate, Cowles Commission for Research in Economics, University of Chicago, Chicago 37, Ill.). **Summarizing a Population of Behavior Patterns.** Abstract. *Econometrica* 16(2):203 Ap '48.*

1956. ASPIN, ALICE A. (Leeds University, Leeds, England). **An Examination and Further Development of a Formula Arising in the Problem of Comparing Two Mean Values.** *Biometrika* 35(1–2):88–96 My '48.*

1957. BACHARACH, A. L. (Glaxo Laboratories Ltd., 983 Greenford Road, Greenford, Middlesex, England). **Biological Assay and Chemical Analysis.** *Analyst* 70(836):394–403 N '45.*

1958. BADGER, E. H. M. (Fulham Laboratory, The Gas Light and Coke Co., Gas Works, Kings Road, London S.W.6, England). **Some Statistical Methods in Refractory Testing.** *Trans Brit Ceramic Soc* 45(1):33–43, discussion 43–4 Ja '46.*

1959. BATEMAN, G. (University College, University of London, London W.C.1, England). **On the Power Function of the Longest Run as a Text for Randomness in a Sequence of Alternatives.** *Biometrika* 35(1–2):97–112 My '48.*

1960. BENTZEL, R., and WOLD, H. **On Statistical Demand Analysis From the Viewpoint of Simultaneous Equations.** *Skandinavisk Aktuarietskrift* 29(1–2):95–114 '46.*

1961. BERGSTRÖM, HARALD. **On the Central Limit Theorem in the Space R_k, $k>1$.** *Skandinavisk Aktuarietskrift* 28(1–2):106–27 '45.*

1962. BOSE, P. K. (Lecturer on Statistics, Calcutta University, Calcutta, India). **Crop Estimation and Its Relation to Agricultural Meteorology.** *Calcutta Stat Assn B* (3):114–9 Mr '48.*

1963. ★BOTT, GEORGE (Operations Evaluation Group (C.N.O.), Navy Department, Washington 25, D. C.). **On the Moments of the Distribution of Autocorrelation.** Master's thesis, Massachusetts Institute of Technology (Cambridge, Mass.), 1947

1964. BROWNLEE, K. A.; DELVES, C. S.; GREEN, C. A.; GRENFELL, E.; JOHNSON, J. D. A.; and SMITH, N. (Research Department, Distillers Co. Ltd., Great Burgh, Epsom, Surrey, England.) **The Biological Assay of Streptomycin by a Modified Cylinder Plate Method.** *J General Microbiol* 2(1):40–53 Ja '48.*

1965. CALLANDAR, W. F. (Assistant Chief in charge of Agricultural Estimates, Bureau of Agricultural Economics, U. S. Department of Agriculture, Washington, D. C.), and SARLE, CHARLES F. **The Bureau of Agricultural Economics Program in Enumerative Sampling.** *J Farm Econ* 29(1):233–6 F '47.*

1966. CAMPBELL, J. A., and EMSLIE, A. R. G. (Vitamin and Physiological Research Laboratory, Central Experimental Farm, Ottawa, Canada.) **Studies on the Chick Assay for Vitamin D: VI, Sources of Variation in the Response of Replicate Groups With Time.** *Poultry Sci* 26(6):568–72 N '47.*

1967. CAMPBELL, J. A., and EMSLIE, A. R. G. **Variability in Chicken Growth Data.** *Poultry Sci* 26(6):573–5 N '47.*

1968. CAMPBELL, J. H. (Electrical Engineer, Engineering Division, Lamp Department, General Electric Co., Nela Park, Cleveland 12, Ohio), and DONSBACH, C. G. (General Electric Co., Schenectady Tube Works, Schenectady, N. Y.). **Analysis of Special Electronic-Tube Tests.** *Trans A.S.M.E.* 68:481–6 Jl '46.*

1969. CARTER, A. H. (King's College, Cambridge, England). **Approximation to Percentage Points of the *z*-Distribution.** *Biometrika* 34(3–4):352–8 D '47.*

1970. CAVALLI, LUIGI (Sezione Genetica Instituto Italiano di Idrobiologia, Pallanza, Novara, Italy), and MAGNI, GIOVANNI (Instituto di Anatomia Patologica Universita, Milan, Italy). **Methods of Analysing the Virulence of Bacteria and Viruses for Genetical Purposes.** *Heredity* 1(1):127–32 Jl '47.*

1971. ★CHAN-MOU, TCHEN. **Mean Value and Correlation Problems Connected With the Motion of Small Particles Suspended in a Turbulent Fluid.** English with Dutch summary. Hague, Netherlands: Martinus Nijhoff (9 Lange Voorhout), 1947. Pp. vii, 125.* *To be reviewed.*

1972. CHURCHILL, H. V. (Chief Chemist, Aluminum Research Laboratories, Aluminum Company of America, New Kensington, Pa.). **A System of Laboratory Evaluation.** *Ind & Eng Chem, Anal Ed* 18(1):66 Ja '46.*

1973. CHURCHILL, H. V., and CHURCHILL, J. R. (Assistant Chief, Aluminum Research Laboratories, Aluminum Company of America, New Kensington, Pa.). **Evaluation of Spectrographic Analytical Data.** *Ind & Eng Chem, Anal Ed* 17(12):751–4 D '45.*

1974. CLAUSEN, JOHN A. (Assistant Professor of Sociology, Cornell University, Ithaca, N. Y.), and FORD, ROBERT N. (American Telephone and Telegraph Co., New York, N. Y.). **Controlling Bias in Mail Questionnaires.** *J Am Stat Assn* 42(240):497–511 D '47.*

1975. COCHRAN, W. G. (Professor of Experimental Statistics, and Director of Research, Institute of Statistics, North Carolina State College, Raleigh, N. C.), and BLISS, C. I. (Biometrician, Connecticut Agricultural Experiment Station, Storrs, Conn.; and Lecturer in Biometrics, Yale University, New Haven, Conn.). **Discriminant Functions With Covariance.** *Ann Math Stat* 19(2):151–76 Je '48.*

1976. COOMBS, CLYDE H. (Assistant Professor of Psychology and Chief of the Research Division, Bureau of Psychological Services, Institute of Human Adjustment, University of Michigan, Ann Arbor, Mich.). **A Rationale for the Measurement of Traits in Individuals.** *Psychometrika* 13(2):59–68 Je '48.*

1977. COOPER, GERSHON (Cowles Commission, University of Chicago, Chicago, Ill.). **Role of Econometric Models in Economic Research.** *J Farm Econ* 30(1):101–16, discussion 139–41 F '48.*

1978. DAVID, F. N., and JOHNSON, N. L. (Lecturers in Statistics, University College, University of London, London W.C.1, England.) **The Probability Integral Transformation When Parameters Are Estimated From the Sample.** *Biometrika* 35(1–2):182–90 My '48.*

1979. DAVIES, O. L.; DUCKWORTH, R. B.; and HARRIS, G. C. M. (Biological Laboratories, Imperial Chemical Industries, Ltd., Hexagon House, Manchester 9, England.) **A Method for Estimating Percentage Germination of Fungal Spores.** Letter. *Nature* 161(4095):642 Ap 24 '48.*

1980. DEGAN, JAMES W. (Psychometric Laboratory, University of Chicago, Chicago 37, Ill.). **A Note on the Effects of Selection in Factor Analysis.** *Psychometrika* 13(2):37–9 Je '48.*

1981. DEMING, W. EDWARDS (Adviser in Sampling, Bureau of the Budget, Washington, D. C.). **On the Classification of Statisticians.** *Am Statistician* 2(2):16–7+ Ap '48.*

1982. DIXON, W. J. (Associate Professor of Mathematics, University of Oregon, Eugene, Ore.), and MOOD, A. M. (Research Statistician, Project RAND, Fourth and Broadway, Santa Monica, Calif.). **A Method for Obtaining and Analyzing Sensitivity Data.** *J Am Stat Assn* 43(241):109–26 Mr '48.*

1983. DUMAS, M. **Sur les courbes de fréquence de K. Pearson. (On the Frequency Curves of K. Pearson).** *Biometrika*

35(1–2):113–7 My '48.*

1984. DURANTE, DOMENICO. Coordination of Principles Useful for Estimating Crop Yields: Quantitative Agroscopy. *Int R Agric* 37(3–4, 5–6):33T–48T, 61T–74T Mr–Ap, My–Je '46.*

1985. DVORETZKY, ARYEH. On Smoothing Through Repeated Averaging: Preliminary Report. Abstract. *B Am Math Soc* 54(3):287–8 Mr '48.*

1986. EDWARDS, ALLEN L. (Professor of Psychology), and KILPATRICK, FRANKLIN P. (The University of Washington, Seattle, Wash.) Scale Analysis and the Measurement of Social Attitudes. *Psychometrika* 13(2):99–114 Je '48.*

1987. EISENHART, CHURCHILL (Chief, Statistical Engineering Laboratory, National Applied Mathematics Laboratories, National Bureau of Standards, Washington, D. C.). The Role of a Statistical Consultant in a Research Organization. *Am Statistician* 2(2):6–7 Ap '48.*

1988. ELDERTON, WILLIAM (President, Equitable Life Assurance Society, 19 Coleman St., London E.C.2, England). Cricket Scores and Some Skew Correlation Distributions. *J Royal Stat Soc* 108(1–2):1–11, discussion 22–39 '45.*

1989. ELLIS, MAX E. (Oscar Mayer & Co.), and RIOPELLE, ARTHUR J. (Research Assistant, Department of Psychology, University of Wisconsin). (Madison, Wis.) An Efficient Punched-Card Method of Computing ΣX, ΣX^2, ΣXY, and Higher Moments. *Psychometrika* 13(2):79–85 Je '48.*

1990. ★ENRICK, NORBERT L. (Associate Professor of Management, Southwestern Louisiana Institute, Lafayette, La.). Quality Control: A Manual of Quality Control Procedure Based Upon Scientific Principles and Simplified for Practical Application in Various Types of Manufacturing Plants. New York 13: Industrial Press (148 Lafayette St.), 1948. Pp. vi, 122. $3.00.* (Brighton 1, England: Machinery Publishing Co., Ltd.) *To be reviewed.*

1991. FELLER, W. (Professor of Mathematics, Cornell University, Ithaca, N. Y.). On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions. *Ann Math Stat* 19(2):177–89 Je '48.*

1992. FINNEY, D. J. (Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford, Oxford, England). The Fisher-Yates Test of Significance in 2×2 Contingency Tables. *Biometrika* 35(1–2):145–56 My '48.*

1993. FINNEY, D. J., and STEVENS, W. L. A Table for the Calculation of the Working Probits and Weights in Probit Analysis. *Biometrika* 35(1–2):191–201 My '48.*

1994. FOG, DAVID. The Geometrical Method in the Theory of Sampling. *Biometrika* 35(1–2):46–54 My '48.*

1995. FRASER, DUNCAN C. (Duncan C. Fraser & Co., Actuaries, 18 Haskins Hey, Liverpool 2, England). A Note on the Gompertz Table. *J Inst Actuaries* 123 (337):423–6 D '47.*

1996. ★FREEMAN, H. A. (Associate Professor of Statistics, Massachusetts Institute of Technology, Cambridge, Mass.); FRIEDMAN, MILTON (Associate Professor of Economics, University of Chicago, Chicago 37, Ill.); MOSTELLER, FREDERICK (Associate Professor of Mathematical Statistics, Department of Social Relations, Harvard University, Cambridge, Mass.); and WALLIS, W. ALLEN (Professor of Statistics and Business Economics, University of Chicago, Chicago 37, Ill.), Editors. Sampling Inspection: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective. By the Statistical Research Group, Columbia University, Applied Mathematics Panel, Office of Scientific Research and Development. New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1948. Pp. xx, 395. $5.25.* *To be reviewed.*

1997. FROE, A DE; HUIZINGA, J.; and GOOL, J. VAN. Variation and Correlationcoefficient. English. *Proc Kon Ned Akad v Wet* 50(7):807–15 S '47.*

1998. FROEHLICH, CLIFFORD P., and KELLER, ROBERT J. (Assistant Professor of Education, and Associate Director, Bureau of Institutional Research, University of Minnesota, Minneapolis, Minn.). A Technique for Obtaining Summary Data From Aggregate Weighting Sheets on the Graphic Item Counter Attachment to the Test Scoring Machine. *Ed & Psychol Meas* 7(3):429–39 Ag '47.*

1999. GAINES, W. L. (Professor of Milk Production; and Chief, Milk Production, Agricultural Experiment Station; University of Illinois, Urbana, Ill.); DAVIS, H. P. (Professor of Dairy Husbandry and Chairman of the Department, University of Nebraska, Lincoln, Neb.); and MORGAN, R. F. (Associate Professor of Journalism and Assistant Professor of Dairy Husbandry, University of Nebraska, Lincoln, Neb.). Within-cow Regression of Milk-Energy Yield on Age and Liveweight.

J Dairy Sci 30(5):273–8 My '47.*

2000. GENGERELLI, J. A. (Associate Professor of Psychology, University of California, Los Angeles, Calif.). A Binomial Method for Analyzing Psychological Functions. *Psychometrika* 13(2):69–77 Je '48.*

2001. GHURYE, S. C. A Characteristic of Species of 7×7 Latin Squares. *Ann Eug* 14(2):133 Ap '48.*

2002. GINI, CORRADO (Director, Institute of Statistics, University of Rome, Rome, Italy). The First Steps of Statistics, pp. 37–45. In *Educational Research Forum Proceedings, Endicott, New York, August 25–29, 1947*. New York 22: International Business Machines Corporation (590 Madison Ave.), 1948. Pp. 96. Paper. Gratis.*

2003. ★GIRSHICK, MEYER A. (Douglas Aircraft Co., Santa Monica, Calif.). Contributions to the Theory of Sequential Analysis. Doctor's thesis, Columbia University (New York, N. Y.), 1947.

2004. GREVILLE, T. N. E. (Principal Actuarial Mathematician, National Office of Vital Statistics, U. S. Public Health Service, Washington, D. C.). Actuarial Note: Adjusted Average Graduation Formulas of Maximum Smoothness. *Rec Am Inst Actuaries* 36(74):249–64 O '47.*

2005. GUILFORD, J. P. (Professor of Psychology, University of Southern California, Los Angeles, Calif.), and MICHAEL, WILLIAM B. (Assistant Professor of Psychology, Princeton University, Princeton, N. J.). Approaches to Univocal Factor Scores. *Psychometrika* 13(1):1–22 Mr '48.*

2006. GURLAND, JOHN (Teaching Assistant in Mathematics, University of California, Berkeley, Calif.). Inversion Formulae for the Distribution of Ratios *Ann Math Stat* 19(2):228–37 Je '48.*

2007. GUTTMAN, LOUIS (Associate Professor of Sociology, Cornell University, Ithaca, N. Y.). An Inequality for Kurtosis. *Ann Math Stat* 19(2):277–8 Je '48.*

2008. GUYE, ROBERTO (Secretary of the Committee on Statistical Education, Inter American Institute, Larrea 1234, Buenos Aires, Argentina). A Program of Action for Promoting Statistical Education in the Western Hemisphere. *Am Statistician* 2(2):10–3 Ap '48.*

2009. HAEMER, KENNETH W. (Statistician, American Telephone and Telegraph Co., 195 Broadway, New York 7, N. Y.). Range-Bar Charts. *Am Statistician* 2(2):23 Ap '48.*

2010. HALDANE, J. B. S. (Weldon Professor of Biometry), and SMITH, CEDRIC A. B. (Assistant Lecturer in Eugenics). (University College, University of London, London W.C.1. England.) A Simple Exact Test for Birth-Order Effect. *Ann Eug* 14(2):117–24 Ap '48.*

2011. HARTLEY, H. O. (Lecturer in Statistics, University College, University of London, London W.C.1, England). The Estimation of Non-Linear Parameters by "Internal Least Squares." *Biometrika* 35(1–2):32–45 My '48.*

2012. HEIDE, J. D. (General Laboratories, United States Rubber Co., Passaic, N. J.). Statistical Control in the Physical Testing of Rubber. *Rubber Chem & Technol* 20(2):577–84 Ap '47.* Reprinted from *India Rubber World* 114(5):653–5+ Ag '46.*

2013. ★HENDRICKS, WALTER A. (Principal Agricultural Statistician, Bureau of Agricultural Economics, Washington 25, D. C.). Mathematics of Sampling. A summary of a course of lectures given during the 1947 Statistical Summer Session at Virginia Polytechnic Institute. Virginia Agricultural Experiment Station, Special Technical Bulletin. Blacksburg 13, Va.: the Station, February 1948. Pp. ii, 45. Gratis.* *To be reviewed.*

2014. HEUX, J. W. N. Le. The Growth-Curve. English. *Proc Kon Ned Akad v Wet* 50(9):1201–13 N '47.*

2015. ★HOMEYER, PAUL G. (Assistant Professor of Statistics and Mathematics); CLEM, MARY A.; and FEDERER, WALTER T. (Resident Collaborator, Bureau of Agricultural Economics, Statistical Laboratory). (Iowa State College, Ames, Iowa.) Punched Card and Calculating Machine Methods for Analyzing Lattice Experiments Including Lattice Squares and the Cubic Lattice. Iowa State College of Agriculture and Mechanic Arts, Agricultural Experiment Station, Statistical Section, Research Bulletin 347. Ames, Iowa: the Station, April 1947. Pp. 25–171. Paper. Gratis.* *To be reviewed.*

2016. HORNER, F. (Radio Division, National Physical Laboratory, Teddington, Middlesex, England). A Problem on the Summation of Simple Harmonic Functions of the Same Amplitude, But of a Random Phase. *Philos Mag* 37(266):145–62 Mr '46.*

2017. HOVANITZ, WILLIAM (Laboratory of Vertebrate Biology, University of Michigan, Ann Arbor, Mich.). A Graphic Method of Illustrating Ecological and Geographical Distributions. *Ecol* 29(1):121 Ja '48.*

2018. HSU, L. C. (Associate Professor of Mathematics, National Tsing-Hua University, Peiping, China). Note on an Asymptotic Expansion of the *n*th Difference of Zero. *Ann Math Stat* 19(2):273–7 Je '48.*

2019. JACKSON, C. H. N. The Analysis of a Tsetse-Fly Population, III. *Ann Eug* 14(2):91–108 Ap '48.*

2020. JARRETT, R. F. (Instructor in Psychology, University of California, Berkeley, Calif.). Per Cent Increase in Output of Selected Personnel as an Index of Test Efficiency. *J Appl Psychol* 32(2):135–45 Ap '48.*

2021. JELONEK, Z. (Polish University College, London, England); FITCH, E. (Formerly Staff for Scientific Research, Signals Research and Development Establishment, Ministry of Supply, Christchurch, Hants, England); and CHALK, J. H. H. (Assistant Lecturer in Mathematics, University College, University of London, London W.C.1, England). Diversity Reception: Statistical Evaluation of Possible Gain. *Wireless Eng* 24(281):54–62 F '47.*

2022. JOHNSON, D. GALE (Department of Economics, University of Chicago, Chicago, Ill.). The Use of Econometric Models in the Study of Agricultural Policy. *J Farm Econ* 30(1):117–30, discussion 139–41 F '48.*

2023. JOHNSON, N. L. (Lecturer in Statistics, University College, University of London, London W.C.1, England). Alternative Systems in the Analysis of Variance. *Biometrika* 35(1–2):80–7 My '48.*

2024. JOHNSON, N. L. Tests of Significance in the Variate Difference Method. *Biometrika* 35(1–2):206–9 My '48.*

2025. JONES, HOWARD L. (Area Supervisor of Results, Accounting Department, Illinois Bell Telephone Co., 309 West Washington St., Chicago 6, Ill.). Exact Lower Moments of Order Statistics in Small Samples From a Normal Distribution. *Ann Math Stat* 19(2):270–3 Je '48.*

2026. JONES, P. C. T., and MOLLISON, J. E. (Soil Microbiology Department, Rothamsted Experimental Station, Harpenden, Herts, England.) A Technique for the Quantitative Estimation of Soil Microorganisms. With a Statistical Note (pp. 65–8) by M. H. QUENOUILLE (Lecturer in Statistics and Head of the Department, Marischal College, Aberdeen University, Aberdeen, Scotland). *J General Microbiol* 2(1):54–69 Ja '48.*

2027. KAC, M. (Associate Professor of Mathematics, Cornell University, Ithaca, N. Y.). On the Characteristic Functions of the Distributions of Estimates of Various Deviations in Samples from a Normal Population. *Ann Math Stat* 19(2):257–61 Je '48.*

2028. ★KELLEY, TRUMAN LEE (Professor of Education, Harvard University, Cambridge, Mass.). The Kelley Statistical Tables, Revised 1948. Cambridge 38, Mass.: Harvard University Press (38 Quincy St.), 1948. Pp. ix, 223. $5.00.* *To be reviewed.*

2029. KEMPTHORNE, O. (Associate Professor of Statistics, Iowa State College, Ames, Iowa). The Factorial Approach to the Weighing Problem. *Ann Math Stat* 19(2):238–45 Je '48.*

2030. KENDALL, D. G. (University Lecturer in Mathematics, Magdalen College, Oxford University, Oxford, England). On Some Modes of Population Growth Leading to R. A. Fisher's Logarithmic Series Distribution. *Biometrika* 35(1–2):6–15 My '48.*

2031. KENDALL, MAURICE G. (Joint Assistant General Manager and Statistician, Chamber of Shipping of the United Kingdom, Bury Court, St. Mary Axe, London E.C.3, England). The Examination of the Royal Statistical Society. *Am Statistician* 2(2):8–9+ Ap '48.*

2032. KINCAID, W. M. (Instructor in Mathematics, University of Michigan, Ann Arbor, Mich.). Solution of Equations by Interpolation. *Ann Math Stat* 19(2):207–19 Je '48.*

2033. KOOPMANS, TJALLING C. (Associate Professor of Economics; Research Associate, Cowles Commission for Research in Economics; University of Chicago, Chicago 37, Ill.). The Identification of Economic Relationships. Abstract. *Econometrica* 16(2):201–3 Ap '48.*

2034. KOSSACK, CARL F. (Associate Professor of Mathematics, Purdue University, Lafayette, Ind.). On the Computation of Zero-Order Correlation Coefficients. *Psychometrika* 13(2):91–3 Je '48.*

2035. LASKEN, HERMAN (11420 Dehn Ave., Inglewood, Calif.). Direct Determination of Compass Settings for Proportionate Area Pie-Charts. *J Am Stat Assn* 43(241):94–5 Mr '48.*

2036. LORD, R. D. (Heriot-Watt College, Edinburgh, Scotland). A Problem on Random Vectors. *Philos Mag* 39(288):66–71 Ja '48.*

2037. LOTKA, ALFRED J. (Assistant Statistician, Metropolitan Life Insurance Co., 1 Madison Ave., New York 10, N. Y.). Application of Recurrent Series in Renewal Theory. *Ann Math Stat* 19(2):190–206 Je '48.*

2038. MATHER, K. (Head of the Genetics Department, John Innes Horticultural Institution); BOWLER, R. G. (Senior Technical Assistant, Medical Unit, London Hospital); CROOKE, A. C. (Senior Assistant, Medical Unit, London Hospital); and MORRIS, C. J. O. R. (Reader in Chemical Pathology, Medical School, London Hospital). (London, England.) **The Precision of the Plasma Volume Determinations by the Evans Blue Method.** *Brit J Exp Pathol* 28(1):12–24 F '47.*

2039. MEIER, NORMAN C. (Director, Bureau of Audience Research, and Associate Professor of Psychology), and BURKE, CLETUS J. (Research Assistant, Department of Psychology). (State University of Iowa, Iowa City, Iowa.) **Laboratory Tests of Sampling Technique.** *Pub Opinion Q* 11(4):586–93 w '47–48.*

2040. MORAN, P. A. P. (Institute of Statistics, Oxford University, Oxford, England). **Rank Correlation and Product-Moment Correlation.** *Biometrika* 35(1–2):203–6 My '48.*

2041. MUHSAM, H. V. (Formerly Department of Statistics, [British] Palestine Government, Jerusalem, Palestine). **Correlation in Growth.** *Human Biol* 19(4):260–9 D '47.*

2042. NAIR, K. R. (Statistics Department, University College, University of London, London W.C.1, England). **The Distribution of the Extreme Deviate From the Sample Mean and Its Studentized Form.** *Biometrika* 35(1–2):118–44 My '48.*

2043. NAIR, K. R. **The Studentized Form of the Extreme Mean Square Test in the Analysis of Variance.** *Biometrika* 35(1–2):16–31 My '48.*

2044. NANDI, H. K. (Calcutta University, Calcutta, India). **A Note on Student's *t* for Paired Samples.** *B Calcutta Math Soc* 39(2):61–4 Je '47.*

2045. NATIONAL CONFERENCE ON UNIFORM TRAFFIC ACCIDENT STATISTICS, COMMITTEE ON USES OF DEVELOPED INFORMATION. **Chap. 9, Basic Accident Analysis Methods,** pp. 153–87. In their *Uses of Traffic Accident Records.* Saugatuck, Conn.: Eno Foundation for Highway Traffic Control, Inc., 1947. Pp. xvii, 187. Paper. Gratis.*

2046. NIESSEN, ABRAHAM M. (Actuary, U. S. Railroad Retirement Board, 844 Rush St., Chicago, Ill.). **Actuarial Estimates for Public Sickness Insurance Plans.** *J Am Stat Assn* 43(241):61–73 Mr '48.*

2047. NORDSIEK, FREDERIC W. (Assistant to the Director, Department of Applied Research, Standard Brands, Inc.) **The Mathematics of Life: Denied the Physical Scientist's Methods of Simplification, the Biologist Makes Increasing Use of Statistics in Analyzing the Results of His Research.** *Technol R* 50(6):313–5+ Ap '48.*

2048. ★PANSE, V. C. **Report on the Scheme for the Improvement of Agricultural Statistics.** Imperial Council of Agricultural Research, [No. 45]. Delhi, India: Manager of Publications, Government of India, 1946. Pp. 68. Paper. *Rs.* 2-4-0 or 3*s.* 9*d.** *To be reviewed.*

2049. PARKER, J. B. (H. M. Nautical Almanac Office, Block E, Ensleigh Hutments, Lansdown, Bath, England). **The Accumulation of Chance Effects and the Gaussian Frequency Distribution.** Letter. *Philos Mag* 38(284):681–2 S '47.*

2050. PATNAIK, P. B. (83 Princes Park Ave., London N.W.11, England). **The Power Function of the Test for the Difference Between Two Proportions in a 2×2 Table.** *Biometrika* 35(1–2):157–75 My '48.*

2051. PERRY, W. L. M. (National Institute for Medical Research, Hampstead, London N.W.3, England). **An Improved Method for the Assay of Toxicity of Arsenicals.** Letter. *Nature* 161(4103):975 Je 19 '48.*

2052. PITTMAN, MARGARET (Bacteriologist, Biologics Control Laboratory, National Institute of Health, U. S. Public Health Service, Bethesda, Md.), and LIEBERMAN, J. E. (Biostatistician, Public Health Methods Division, U. S. Public Health Service, Washington, D. C.). **An Analysis of the Wilson-Worcester Method for Determining the Median Effective Dose of Pertussis Vaccine.** *Am J Pub Health* 38(1, pt 1):15–21 Ja '48.*

2053. PRAKASH, OM. (Seth G. B. Podar College, Nawalgarh, Jaipur, India). **The Mean Deviation in Grouped Data.** Letter. *J Inst Actuaries Students' Soc* 7(1):52–3 Jl '47.*

2054. PRESTON, FRANK W. **Quality Control Techniques Applicable to Glass Tank Operation.** *Glass Ind* 26(10):465–6 O '45.*

2055. RAINWATER, L. J. (Assistant Professor of Physics), and WU, C. S. (Pupin Physics Laboratory). (Columbia University, New York 27, N. Y.) **Application of Probability Theory to Nuclear Particle Detection.** *Nucleonics* 1(2):60–9 O '47; 2(1):42–9 Ja '48.*

2056. RAO, C. RADHAKRISHNA (King's College, University of Cambridge,

Cambridge, England). **Tests of Significance in Multivariate Analysis.** *Biometrika* 35(1–2):58–79 My '48.*

2057. RASCH, G. (State Serum Institute; and Assistant Professor, University of Copenhagen; Copenhagen, Denmark). **A Functional Equation for Wishart's Distribution.** *Ann Math Stat* 19(2):262–6 Je '48.*

2058. REITZ, L. K., and SILLAY, F. J. (Paper Service Department, Eastman Kodak Co., Rochester, N. Y.) **Application of Statistical Methods to Paper Testing Procedures: Part I, Folding Endurance.** *Paper Trade J* 126(17):54–8 Ap 22 '48.* Abstract. *Paper Trade J* 126(8):54 F 19 '48.*

2059. RENNER, MAYNARD S. **Some Basic Concepts of Statistics and Their Application to Simple Industrial Experimentation.** Abstract. *Paper Trade J* 126(8):55 F 19 '48.*

2060. ROBBINS, HERBERT (Associate Professor of Mathematical Statistics, University of North Carolina, Chapel Hill, N. C.). **The Distribution of a Definite Quadratic Form.** *Ann Math Stat* 19(2):266–70 Je '48.*

2061. ROBBINS, HERBERT. **On the Asymptotic Distribution of the Sum of a Random Number of Random Variables.** *Proc Nat Acad Sci* 34(4):162–3 Ap '48.*

2062. ROOS, CHARLES F. (President, Econometric Institute, Inc., 500 Fifth Ave., New York 18, N. Y.). **A Future Role for the Econometric Society in International Statistics.** *Econometrica* 16(2):127–34 Ap '48.*

2063. ★RUTHERFORD, JOHN G. (Quality Control Methods and Planning, Glenn L. Martin Co., Baltimore, Md.). **Quality Control in Industry: Methods and Systems.** New York 19: Pitman Publishing Corp. (2 West 45th St.), 1948. Pp. xvii, 201. $3.50.* *To be reviewed.*

2064. SAMUELSON, PAUL A. (Professor of Economics, Massachusetts Institute of Technology, Cambridge, Mass.). **Exact Distribution of Continuous Variables in Sequential Analysis.** *Econometrica* 16(2):191–8 Ap '48.*

2065. SARLE, CHARLES F. (Head, Division of Special Farm Statistics, Bureau of Agricultural Economics, U. S. Department of Agriculture, Washington, D. C.). **The Place for Enumerative Surveys in Research and Graduate Training in Agricultural Economics.** *J Farm Econ* 29(1): 237–40 F '47.*

2066. SCATES, DOUGLAS E. (Professor of Education, Duke University, Durham, N. C.) **There's a Place for Statistics in General Education.** *Nations Sch* 41(2):42–3 F '48.*

2067. SCATES, DOUGLAS E., and DEMING, W. EDWARDS. **Education in Statistics for Participation in Current Affairs.** *Sch R* 56(5):262–9 My '48.*

2068. SCHAEFER, WILLIS C. (Assistant Professor of Psychology, University of Maryland, College Park, Md.). **A Simplified Procedure for Item Analysis of Examination Material,** pp. 64–70. In *Educational Research Forum Proceedings, Endicott, New York, August 25–29, 1947.* New York 22: International Business Machines Corporation (590 Madison Ave.), 1948. Pp. 96. Paper. Gratis.*

2069. SCHLOSBERG, HAROLD (Associate Professor of Psychology, Brown University, Providence, R. I.). **A Probability Formulation of the Hunter-Sigler Effect.** *J Exp Psychol* 38(2):155–67 Ap '48.*

2070. SCHMID, JOHN, JR. (Instructor in Mathematics, University of Wisconsin High School. Madison, Wis.). **The Relationship Between the Coefficient of Correlation and the Angle Included Between Regression Lines.** *J Ed Res* 41(4):311–3 D '47.*

2071. SCHÜTZENBERGER, MARCO P. (21 rue de Malte, Paris XI, France). **An Abac for the Sample Range.** *Psychometrika* 13(2):95–7 Je '48.*

2072. SEAL, H. L. (Actuary, National Life Insurance Co., Toronto, Canada) **A Note on the χ^2 Smooth Test.** *Biometrika* 35(1–2):202 My '48.*

2073. SETH, G. R. (2 Imperial Ave. Old Viceregal Estate, Delhi, India). **The Variance of Sequential Estimates.** Abstract. *B Am Math Soc* 54(3):288 Mr '48.*

2074. SICHEL, HERBERT S. **An Experimental and Theoretical Investigation of Bias Error in Mine Sampling With Special Reference to Narrow Gold Reefs.** Discussion. *B Inst Mining & Metallurgy* (485, 486) 26–41, 55–9 Ap, My '47.*

2075. SILBER, JACK (Associate Professor of Mathematics, Roosevelt College, 430 South Michigan Ave., Chicago 5, Ill.). **Multiple Sampling for Variables.** *Ann Math Stat* 19(2):246–56 Je '48.*

2076. SMIRNOV, N. **Table for Estimating the Goodness of Fit of Empirical Distributions.** *Ann Math Stat* 19(2):279–81 Je '48.*

2077. SMITH, H. FAIRFIELD (Rubber Institute of Malaya, P.O. Box 150, Kuala Lampur, Federated Malay States). **Standard Errors of Means in Sampling Surveys With Two-Stage Sampling.** *J*

Royal Stat Soc 110(3):257–9 pt 3 '47.*

2078. SNEDECOR, GEORGE W. (Professor of Statistics, Iowa State College, Ames, Iowa). A Proposed Basic Course in Statistics. *J Am Stat Assn* 43(241):53–60 Mr '48.*

2079. SNYDER, THOMAS L. (Bacteriologist, MV Division, Research and Development Department, Camp Detrick, Frederick, Md.). The Relative Errors of Bacteriological Plate Counting Methods. *J Bacteriol* 54(5):641–54 N '47.*

2080. SOLOMON, L. (Statistical Adviser to the Director of Naval Air Organization and Training, Naval Staff, Admiralty, London S.W.1, England). Statistical Estimation. *J Inst Actuaries Students' Soc* 7(3, 4):144–73, 213–34 Ja, Ap '48.*

2081. SPIELMANS, JOHN V. (Associate Professor of Economics, Marquette University, Milwaukee 3, Wis.). Profile Graphs. *J Am Stat Assn* 43(241):96–108 Mr '48.*

2082. ★SPRINGER, MELVIN D. (Instructor in Mathematics, University of Illinois, Urbana, Ill.). Joint Sampling Distribution of Mean and Standard Deviation for a Chi-Square Universe. Doctor's thesis, University of Illinois (Urbana, Ill.), 1947. Pp. 70.

2083. SPURR, WILLIAM A. (Professor of Business Statistics), and ARNOLD, DAVID R. (Gas Engineer, Superior Oil Co., 1501 Fourth Ave., Los Angeles 6, Calif.). A Short-Cut Method of Fitting a Logistic Curve. *J Am Stat Assn* 43(241):127–34 Mr '48.*

2084. STEFFENSEN, J. F. On Certain Formulas of Mechanical Quadrature. *Skandinavisk Aktuarietskrift* 28(1–2):1–19 '45.*

2085. STEPHAN, FREDERICK F. (Director of Research Program in Education and Professor of Sociology, Princeton University, Princeton, N. J.). History of the Uses of Modern Sampling Procedures. *J Am Stat Assn* 43(241):12–39 Mr '48.*

2086. STEVENS, NEIL E. (Professor of Plant Pathology, University of Illinois, Urbana, Ill.). The Anecdote as an Antidote to Statistical Analysis. *Chronica Botanica* 11(3):188–90 '47.*

2087. STOCK, J. STEVENS (Eastern Service Representative), and HOCHSTIM, JOSEPH R. (Chief Statistician). (Opinion Research Corporation, 44 Nassau St., Princeton, N. J.). Commercial Uses of Sampling. Abstract. *Econometrica* 16(1):76–7 Ja '48.*

2088. STONE, RICHARD (Director, Department of Applied Economics, King's College, Cambridge University, Cambridge, England). On the Interdependence of Blocks of Transactions. *Sup J Royal Stat Soc* 9(1–2):1–32, discussion 32–45 '47.*

2089. STROHKARCK, FRANK (Statistician), and PHELPS, KATHERINE (Assistant Statistician). (Research Department, Curtis Publishing Co., Independence Square, Philadelphia 5, Pa.) The Mechanics of Constructing a Market Area Map. *J Marketing* 12(4):493–6 Ap '48.*

2090. ★STUBBINGS, G. W. Engineering Statistics: Basic Ideas on the Subject as It Is Applied in Production Control and Standardisation. Mechanical World Monographs, No. 38. Manchester 3, England: Emmott & Co. Ltd., 1947. Pp. 24. Paper. 2*s*.*

2091. SWINEFORD, FRANCES (Research Associate, Department of Education, University of Chicago, Chicago, Ill.). A Table for Estimating the Significance of the Difference Between Correlated Percentages. *Psychometrika* 13(1):23–5 Mr '48.*

2092. SWINEFORD, FRANCES, and HOLZINGER, KARL J. (Professor of Education, University of Chicago, Chicago, Ill.). Selected References on Statistics, the Theory of Test Construction, and Factor Analysis. *Sch R* 56(6):361–6 Je '48.*

2093. SYMES, N. G. (Personnel Officer, Marks & Spencer Ltd., 82 Baker St., London N.W.1, England). Statistics and Personnel Management. *J Inst Personnel Mgmt* 29(293):212–8 S–O '47.*

2094. TAYLOR, E. K. (Chief, Performance Evaluation and Criterion Research Unit), and GAYLORD, R. H. (Research Psychologist). (Personnel Research Section, Adjutant General's Office, Department of the Army, Washington, D. C.) Table for Use in the Computation of Statistics of Dichotomous and Truncated Distributions. *Ed & Psychol Meas* 7(3):441–56 Ag '47.*

2095. THOMPSON, W. R. (Senior Biochemist, Division of Laboratories and Research, New York State Department of Health, Albany, N. Y.). Direct Probability Sequential Analysis: Preliminary Report. Abstract. *B Am Math Soc* 54(3):288–9 Mr '48.*

2096. UVEN, M. J. VAN. Extension of Pearson's Probability Distributions to Two Variables. English. *Proc Kon Ned Akad v Wet* 50(9), 10):1063–70, 1252–64 N, D '47; 51(1, 2):41–52, 191–6 Ja, F '48.*

2097. VAJDA, S. (Statistician, Statistical Branch, Admiralty, London S.W.1, England). An Outline of the Theory of the 'Analysis of Variance.' *J Inst Actuaries*

Students' Soc 7(4):235–52 Ap '48.*

2098. VALLE, J. D. DALLA (Consulting Engineer, New York, N. Y.). The Probability Theory Applied to Exhaust System Design. *Heat & Vent* 44(7):95–6 Jl '47.*

2099. WAINWRIGHT, C. (Chief Technical Officer, Bakelite, Ltd., 18 Grosvenor Gardens, London S.W.1, England). Statistical Quality Control: Elementary Principles in the Preparation of Specifications for Plastics. *Brit Plastics* 18(211):524–30 D '46.*

2100. WALD, ABRAHAM (Professor of Mathematical Statistics and Executive Officer of the Department, Columbia University, New York, N. Y.). Estimation of a Parameter When the Number of Unknown Parameters Increases Indefinitely With the Number of Observations. *Ann Math Stat* 19(2):220–7 Je '48.*

2101. WALD, ABRAHAM. On the Distribution of the Maximum of Successive Cumulative Sums of Independently But Not Identically Distributed Chance Variables. *B Am Math Soc* 54(4):422–30 Ap '48.*

2102. WALKER, HELEN M. (Professor of Education, Teachers College, Columbia University, New York 27, N. Y.). Statistical Understandings Every Teacher Needs. *Teach Col Rec* 49(7):452–7 Ap '48.*

2103. WALLACE, W. N. W. (*Formerly* Quality Control Engineer, Phillips Electrical Industries Pty. Ltd., Alberton, South Australia). Control and Chemical Analysis. *Austral Chem Inst J & Proc* 12(7):243–8 Jl '45.*

2104. WALLACE. W. N. W. Quality Control (Statistical) as an Aid to Analytical Control. *Austral Chem Inst J & Proc* 12(7):239–42 Jl '45.*

2105. WEISS, EDWARDS S. (Statistician, Tuberculosis Control Division, U. S. Public Health Service, Washington, D. C.). An Abridged Table of Probits for Use in the Graphic Solution of the Dosage-Effect Curve. *Am J Pub Health* 38(1, pt 1):22–4 Ja '48.*

2106. WESTENBERG, J. Mathematics of Pollen Diagrams. English. *Proc Kon Ned Akad v Wet* 50(5, 6):509–20, 640–8 My, Je '47.*

2107. WILLIAMS, KENNEDY (Assistant Actuary, Friends Provident & Century, Life Office, 7 Leadenhall St., London E.C.3, England). The Classical Theory of Risk—A Statistical Approach. *J Inst Actuaries Students' Soc* 7(3):126–43 Ja '48.*

2108. WISHART, JOHN (Reader in Statistics, University of Cambridge, Cambridge, England). Proofs of the Distribution Law of the Second Order Moment Statistics. *Biometrika* 35(1–2):55–7 My '48.*

2109. WOLD, HERMAN (Director, Institute of Statistics, University of Uppsala, Uppsala, Sweden). A Theorem on Regression Coefficients Obtained From Successively Extended Sets of Variables. *Skandinavisk Aktuarietskrift* 28(3–4): 181–200 '45.*

2110. WOOD, GEORGE H. Cricket Scores and Geometrical Progression. *J Royal Stat Soc* 108(1–2):12–22, discussion 22–39 '45.*

2111. WOOLF, B. (Usher Institute of Public Health, Warrender Park Road, Edinburgh 9, Scotland). Economy in Cell Counts. Letter. *Nature* 161(4088):361 Mr 6 '48.*

2112. ★WORSLEY, BEATRICE HELEN (Research Associate, Computational Project, Department of Applied Mathematics, University of Toronto, Ontario, Canada). Mathematical Survey of Calculating Devices, With an Appendix on an Error Analysis of Differential Analysers. Master's thesis, Massachusetts Institute of Technology (Cambridge, Mass.), 1947.

2113. WOYTINSKY, W. S. (Research Director, Employment and Wage Survey, Twentieth Century Fund, 1302 18th St., N.W., Washington 6, D. C.). Consumption-Saving Function: Its Algebra and Philosophy. *R Econ & Stat* 30(1):45–55 F '48.*

2114. YATES, F. (Head, Statistical Department, Rothamsted Experimental Station, Harpenden, Herts, England). The Analysis of Contingency Tables With Groupings Based on Quantitative Characters. *Biometrika* 35(1–2):176–81 My '48.*

2115. ★YOUNG, CHENG-PANG. On Certain Methods of Testing Statistical Hypotheses and Applications. Master's thesis, George Washington University (Washington, D. C.), 1946.

2116. ZUCKER, LOIS M. (Research Associate, Department of Pathology, College of Physicians and Surgeons, Columbia University, New York 32, N. Y.). Evaluation of Slopes and Intercepts of Straight Lines. *Human Biol* 19(4):231–59 D '47.*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Volume 43 December 1948 Number 244

ARTICLES

Commercial Uses of Sampling . . . . J. Stevens Stock and Joseph R. Hochstim 509

Variation of the Frequency of Fatal Quarrels with Magnitude . . . . Lewis F. Richardson 523

Bank Reserves and Business Fluctuations. . . . Clark Warburton 547

The Ordering of n Items Assigned to k Rank Categories by Votes of m Individuals . . . . Garret L. Schuyler 559

Levels of Significance for Variance Ratio of Two Samples of Equal Size . . . . C. J. Kirchen 564

Correction to "On Estimating Precision of Measuring Instruments and Product Variability" . . . . Frank E. Grubbs 564

Main Effects and Interactions . . . . D. J. Finney 566

A Test for Symmetry in Contingency Tables . . Albert H. Bowker 572

The War Production Board's Statistical Reporting Experience, Part IV . . . . David Novick and George A. Steiner 575

Statistical Methodology Index . . . . Oscar Krisen Buros 616

Index of *Journal*, Volume 43, 1948 (Nos. 241, 242, 243, 244) 619

Articles, by author 620

Book Reviews, by author 621

BOOK REVIEWS

Institute of Mathematical Statistics, *The Teaching of Statistics* . . . . Alan E. Treloar 597

Royal Statistical Society, Committee on the Teaching of Statistics, *The Teaching of Statistics in Universities and Colleges* . Alan E. Treloar 597

Florence, P. Sargant, *Investment, Location, and Size of Plant: A Realistic Inquiry into the Structure of British and American Industries* . . . . Edgar M. Hoover 601

Samuelson, Paul Anthony, *Foundations of Economic Analysis* . . . . George J. Stigler 603

Columbia University Statistical Research Group, Applied Mathematics Panel, Office of Scientific Research and Development, Edited by H. A. Freeman, *Sampling Inspection: Principles, Procedures and Tables for Single, Double and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective* . . Edwin G. Olds 605
. . . . Paul S. Olmstead 611

Weld, Walter E., *How to Chart: Facts from Figures with Graphs* . . . . Kenneth W. Haemer 613

Index to Volumes 1–34, 1888–1939, may be obtained from the ASA. The Journal is also indexed in the Industrial Arts Index, the International Index to Periodicals, and the Public Affairs Information Service Bulletin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 244 DECEMBER 1948 *Volume 43*

COMMERCIAL USES OF SAMPLING*

J. STEVENS STOCK AND JOSEPH R. HOCHSTIM
Opinion Research Corporation, Princeton, New Jersey

To illustrate how a commercial research agency meets its sampling problems, eight surveys of widely varying types are briefly described.

I

THE PURPOSE of this paper is to report on some of the commercial uses of sampling. We shall describe a series of actual examples which show how a commercial research agency meets its sampling problems. As in all sampling problems, the nature of the research objective and the facilities at hand are the all-important conditions which determine the approach.

1. Most commercial agencies must be prepared to tackle an extremely wide variety of sampling problems. No one sampling procedure or principle is applicable in all, or even in a majority of, cases.
2. The actual budget is frequently not always commensurate with the difficulty of the problem. Often it is necessary to do the best job possible under a very restrictive budget, even though a larger budget would allow for much greater dollar efficiency.
3. Most commercial research problems are geared into current-policy determination. Consequently, research efficiency must sometimes be sacrificed for speed.
4. Although there are important exceptions, commercial research problems are for the most part nonrecurrent. Time series and repetitive studies are conducted, but are relatively small in number.
5. Many research problems must be tackled with limited source material because of the specialized nature of the problems.
6. Many commercial research studies must cover unusual, difficult-to-define universes of people, such as the purchasers of a certain commodity, all chemists in the United States, all people who will vote in the next election, etc.

---

* Presented at the 25th Session of the International Statistical Institute, a constituent part of the International Statistical Conferences, Washington, D. C, September 6–18, 1947. This paper will also appear in the Proceedings of the Conferences, which will be published in the near future.

7. Commercial research agencies often find it necessary to study attitudes and opinions which have highly emotional, personal, or moral implications in the respondent's mind.
8. Very often, even in large national studies, the commercial research agency must make detailed analyses for relatively small areas within the country. This makes it necessary to maintain extremely large part-time staffs widely distributed throughout the nation.

Regardless of any pet theories, the sampler has to consider the following:

1. The problems the user of the data needs to have answered.
2. The actions likely to be based on the results of the survey.
3. The degree of precision required.
4. The size of the budget.

In any case, the sample design chosen should provide the user of the data with maximum information at minimum cost.

II

Our first example involves the sampling of purchasers of alcoholic beverages. It is illustrative of the quick, nonrecurrent survey of a difficult-to-define universe on a subject that for some respondents has moral, personal, and political implications. In one particular state a law had been passed increasing the liquor tax. The client, a distilling company, wanted to know who was getting the blame from purchasers of hard liquor for the increase in prices—the state government, the distillers, or the distributors. The answer to this question would determine the direction the client's public-relations effort would take and the issues on which it would concentrate.

Early testing of the questionnaire indicated that people were unable or unwilling to give accurate answers as to their purchases of alcoholic beverages, except at the time of purchase. The client, however, was interested in directing his public-relations effort at the purchasers of his product so that it would have the greatest impact. It was necessary, therefore, to design the sample so that the number of respondents was approximately proportionate to the amount of total liquor sales throughout the universe.

In the state being surveyed packaged alcoholic beverages are sold only in state liquor stores. Data on liquor sales were available for the past year for different districts of the state, for individual towns in each county, and for individual stores. The sample was designed as follows:

The number of interviews with purchasers in the various districts was computed in proportion to the amount of liquor sales in each district. The two largest cities of the state were removed from their districts and treated as separate units. Interviews were allocated to these two cities separately according to the amount of their own liquor sales.

Each district was divided into three groups of counties: those with high per-capita consumption, those with medium, and those with low per-capita consumption.

Within each of these three groups of counties the towns were stratified by per-capita volume of liquor sales. One sample town was selected from each substratum. The number of interviews with liquor purchasers allocated to each sample town was in proportion to the total volume of liquor sales of the substratum from which the town was selected. Within each town a list of stores was prepared and arranged in rank order according to volume of liquor sales. Stores were then selected within each sample town systematically with probability of selection proportionate to sales volume.

Interviewers were instructed to stand outside the stores selected and to interview every *n*th person leaving the stores until their allocated number of interviews was complete. In order to assure proper distribution by time of day and day of the week, interviewers were instructed to interview half of their respondents on Friday afternoon and Saturday, and the other half on other days of the week. In addition, the interviewers were instructed to fill half their their quotas before 4:00 P.M. each day and half of them after 4:00 P.M. each day.

III

Our next example is typical of a study of the community relations of an industrial firm.

The largest employer in a Midwestern town of about 100,000 population felt that his company had done a good job of public and industrial relations. But in spite of all its efforts the company had a strike. The executives wanted to know very quickly what could be done to correct an apparent hostility of the citizenry of the community. The problem of the survey was to find what the people in the community were thinking about the company and its employees and what the company could do in the opinion of the community to make the plant a better place to work.

The 1940 Census Block Statistics, which were available for this city, were brought up to date in numbers of dwelling units with the aid of the city planner's data on new construction. All blocks were then stratified first by census tract, then by average rent. Within each of these strata they were arranged by the number of dwelling units per block. The sample blocks were selected by systematic random sampling with probability of inclusion proportionate to the number of dwelling units in each block.

Within each block the interviewer was given a starting point and instructed to go in clockwise direction and to interview every *n*th dwelling unit in the assigned block until the quota for that block was filled.

In cases of noninterview because of refusal or because no one was at home, interviewers were told to go to the next dwelling unit, but to continue the count from the dwelling unit originally designated. In order to minimize the bias resulting from the omission of not-at-homes and to give more adequate opportunity for getting employed workers in the sample, interviewers were required to do their interviewing in the evening and over weekends.

Quotas covering the whole city by sex and age were given to the supervisor. By cumulative daily tallies of the interviews returned, the supervisor was able to watch the representativeness of the survey. He assigned varying quotas to the interviewers for the selection of individuals within the sample households when certain age-sex groups in the population appeared to be under-represented according to the over-all quota.

Note: Recent experience with this combination of areal and quota techniques indicates that much of the interviewer's tendency toward bias is eliminated by the systematic pre-selection of each dwelling unit in which an interview is to take place. The use of quotas for the selection of an individual within each household allows the interviewer very little free choice.

IV

The next example is a survey to measure the attitudes of social-science teachers toward basic principles of free-enterprise economics. The particularly difficult problem in this case was to get a cross section of a highly specialized population, namely, college and high-school teachers of government, economics, sociology, and history.

The sample of high-school teachers was allotted by states according to data on high-school enrollment published by the Office of Education. Interviews were allocated to cities and towns in proportion to the population distribution of urban places within each state. The selection of the 160 cities and towns within states and size-of-town groups was determined by the location of the company's resident interviewers.

Note: At first glance, it would seem the selection of towns to suit the convenience of the resident interviewers is quite unrigorous. In this case, however, it is an excellent means of selecting a random sample of towns. Since there is no conceivable correlation between the methods by which part-time resident interviewers are recruited and any attitude characteristic of the social-science teachers being studied, we can say that we have selected our sample at random.

It is a common practice in commercial research firms (and others as well, we presume) to base the selection of localities for the cross section on the location of the resident staff. Many suspect this technique of bias. However, as long as the investigator can guarantee that the method of recruiting is in no predictable way connected with the problem studied, we can be assured that the selection of localities is as random as if it had been based on the use of random numbers or some game of chance.

Most commercial research agencies that maintain large part-time resident interviewing staffs recruit them in proportion to the distribution of the population by region, state, and size of town groups within states. In this case the total maximum number of resident interviewers available was approximately 1,000 located in more than 600 places.

As for college teachers, a complete list of all colleges was prepared by type—senior, junior and teacher colleges—and was arranged by states. The number of teachers was recorded for each college, and the colleges were put into two groups: those with a teaching staff of 200 or more and those with fewer than 200 teachers. Every *n*th college was selected for the sample.

Since in neither case was there available a total count on any social or economic characteristic of teachers of social-science subjects, it was impossible to assign any quotas by characteristics. It was necessary therefore, to insure a strict randomization of the respondents. Accordingly, interviewers were instructed to obtain a complete list of social-science teachers for the sample colleges and all high schools in the towns selected. The interviewers were further asked to alphabetize this list and to interview the first *n* teachers starting with a given letter of the alphabet. Interviewers were given different starting letters.

A separate part of this survey illustrates the use of mail questionnaires in studying a highly specialized universe. A large industrial company had conducted for social-science teachers a course of several weeks in some of the social problems faced by business and the techniques used in solving these problems. The purpose of this study was to measure the impact on teachers of impressions obtained in this course.

A complete list of all participants in the course was obtained. Post cards were mailed to participants asking them to state by return mail whether they would be willing to answer a mail questionnaire.

A great majority of the teachers expressed their willingness to answer the mail questionnaires. When the returns were counted, however, slightly less than 50% of all participants actually mailed questionnaires.

Since it was felt that the returned sample might not be representative and that those teachers who had not returned the mail questionnaires might be different in their characteristics and attitudes, it was decided to take a separate sample of nonreturners by personal interviews.

The industrial company that gave the course was located in a very large Midwestern city, and about 40% of the teachers attending the course lived within an area of 15 miles from the plant city. It was decided that resident interviewers were to telephone all social-science teachers included in the original list who lived within a 15-mile radius of the plant city. They were to inquire whether the teachers had returned the mail questionnaires and if not, to make an appointment for a personal interview.

By this method, interviews with approximately one fourth of the teachers who did not answer the mail questionnaires were secured. It was then assumed that those who did not answer the questionnaires and who lived in the plant-city area were representative of all the teachers not answering the questionnaire. Accordingly, the answers to personal interviews were weighted to give proportionate representation to those who had, and those who had not, answered the mail questionnaires.

V

Much commercial research is devoted to studying the style or product preferences of highly specialized groups in the population who are actual or potential customers for a certain commodity. The purpose of such research is to get rough estimates of relative customer preference in order to guide manufacturing and sales policy.

A survey of this kind was recently conducted to determine style preferences for a particular piece of furniture. In order to set up as nearly as possible a realistic buying situation, the study was restricted to three separate groups:

1. Well-to-do respondents were questioned on styles in high-priced items.

2. Average and medium-well-to-do respondents were questioned on medium-priced items.
3. Poor respondents were questioned on styles in low priced items.

To simplify the approach and meet the restrictive budget, interviews were conducted only in towns with population of 50,000 or more. All towns having 50,000 or more population were stratified by geographic sections. Interviews were allocated to cities and towns in proportion to their population in groups of 15 or multiples of 15. In the case of small cities, this amounted to arraying them by population size and selecting an ordered random sample of cities. However, instead of selecting every *n*th city, we selected every *n*th person, so that the probability of selecting any one city was determined by the chance of selecting any one person living in the city. Therefore, the probability of selecting each city was proportionate to its population size.

Note: A common practice of commercial research agencies is to assign interviews in groups of 10 or 15, in order to give each interviewer an assignment large enough to make it worth his while to work on the study and small enough so that he can finish his interviewing in the time limits set for the study.

Within each of the cities selected a relatively large primary sample of 100 to 300 blocks was chosen by selecting every *n*th block from the published block statistics. From these large samples of blocks smaller samples of blocks were selected for assignment.

Note: This practice of selecting relatively large primary samples of blocks, then stratifying them by section of the city, rent, and numbers of dwelling units, for further selection of smaller samples of blocks amounts to *double sampling.* It appears to be a very useful practice where small samples of blocks from relatively large cities are needed. The very practical reason for this kind of double sampling is to save on the burdensome clerical time that would be required to stratify all the thousands and tens of thousands of blocks in a large city. It is necessary only to stratify the primary sample of 100 to 300 blocks.

Within each city the primary sample of blocks was divided into three broad groups: high-rent, medium-rent, and low-rent blocks. In the large cities in the sample, where as many as three groups of 15 interviews were assigned, three interviewers were assigned to sample blocks corresponding to each of the three rental groups. In smaller cities, where only one or two sets of 15 interviews were assigned, interviewers were required to interview in sample blocks from (correspondingly) only one or two of the rental types. Thus, no attempt was made in smaller cities to get a cross section of each city since no individual city analyses were required.

Note: The reason for this assignment procedure was that a different questionnaire was used for each type of respondent—the well-to-do, the average, and the poor. Because of the complicated nature of the interview and the danger of confusing the questionnaires, only one type of questionnaire was assigned to each interviewer.

Blocks within each of the three broad rental groups in each sample city were selected by stratifying the primary sample of blocks for each broad rental group into three strata according to rent. From each of these three rental substrata, three blocks were selected at random with probability of selection in proportion to the number of dwelling units on each block. Thus, each interviewer was assigned one type of questionnaire in one broad rental group. Further, he was assigned three blocks from each of his three rental substrata (9 blocks in all). The interviewer was instructed to take five interviews from each set of the three blocks in each rental substratum. He was also instructed to go around as many of the three blocks as necessary to contact respondents owning the particular piece of furniture or saying they were in the market for it. If the respondent did not meet this qualification, the interview was discontinued and contact was made with the next dwelling unit. Wherever the respondent owned such a piece of furniture or was in the market for one, the interview was completed.

Various styles of furniture of different price range were judged by respondents (higher-priced pieces by respondents living in the higher-rent blocks, lower-priced pieces by respondents in the lower-rent blocks). On a few pieces of medium price, however, we measured preferences in all three groups for purposes of experimental control.

VI

A frequent problem for any research agency is the high expense of travel from interview to interview. The next example illustrates a problem in studying farmers' opinions, one of the worst situations from the standpoint of interviewer travel cost.

During war time, government regulations determined prices and conditions under which much farm produce could be sold by farmers to food manufacturers and processers. There was considerable resentment on the part of the farmers against these regulations. There was also resentment against the manufacturers who in the opinion of some farmers manipulated these regulations to their own advantage and to the farmer's disadvantage.

The purpose of the survey was to study farmers' sentiments on this issue. It was decided to restrict the study to seven counties in a Midwestern state—an area where the client did most of his purchasing.

Interviews were allotted to each county in proportion to number of farms, but in order to save cost, interviewing was limited to every other township.

State highway maps were obtained on which square-mile sections within each sample township were numbered.

The number of sections within each sample township was recorded, and a cumulative total of these sections was obtained. Then the average number of farms per section was computed. In each county the desired number of sections was obtained by dividing the assigned number of interviews by the average number of farms per section. In order to concentrate interviewing still more, two adjacent sections were combined into one double section. The sampling ratio was then obtained by dividing the total number of sections in the sample of townships by the number of double sections wanted. An ordered sample of every *n*th double section was selected. Interviewers were instructed to interview every farmer in the double sections selected, making call-backs wherever the farmer was not found at home.

Note: This concentration of interviews in double sections is, of course less efficient in a statistical sense than a sample of single sections or individual farms. In the information-per-dollar sense, however, it is undoubtedly much more efficient. Moreover, it greatly facilitates call-backs.

VII

Frequently firms desire to make opinion studies among their customers, or associations want to poll a cross section of their members. Often in such cases the client firm or association can supply a complete list of the universe to be studied. The sampling problem is then usually very easy; we stratify the names on the list as efficiently as possible and select every *n*th name.

Shortly before the end of the war, the planning board of a metropolitan area was concerned with the potential post-war employment in the area. During war time the population of the area had increased very considerably, and it was important to make an estimate of the number of people for whom employment might be provided in the years to come. The budget for the survey allowed interviews with 1,500 employers in the area. A complete set of cards showing each firm, its industry group, and the numbers of its employees was made available by the client. In the area to be studied there were approximately 750 firms having 50 employees or more, which accounted for about 54% of all employees in the region. It was decided to include all these 750 firms in the sample.

Of the remaining 15,000 firms having fewer than 50 employees every 20th one was selected.

Thus, approximately equal numbers of firms having more and fewer than 50 employees were selected for the sample. By full coverage of firms that were more important in terms of employment, the statistical error in projecting the average number of employees was reduced considerably.

Note: This example illustrates the common case in which certain segments of the population to be studied contribute such high variability to the final statistical estimate that it is efficient to sample disproportionately from different strata. This amounts to a very simple application of what is commonly called "Neyman sampling."

The cards for employers of 50 persons or less were stratified by detailed industry groups, and were arranged within each group in ascending order from the smallest to the largest firms. Every 20th card of each industry group was then selected, carrying over the count from one industry group to the next.

VIII

The next survey, which uses quota sampling, illustrates a very important and useful technique for studying many different and widespread populations at the same time.

In this instance the purpose of the survey was to bring out pertinent facts bearing upon the public-relations program of a particular industry. A variety of issues was to be covered, and each issue was to be explored thoroughly.

The research problem was to describe in detail the attitudes of many different "publics" toward this industry:

1. Manual workers employed in this particular industry.
2. Residents of each of the major areas (each was analyzed separately) in which the industry is concentrated.
3. Separate cross sections of the general population in seven geographic regions.
4. A national cross section of the continental United States.

Note: More and more in public-relations research it has been found useful to study the "concentric circles" of populations toward which the public-relations effort is being directed. This puts great premium on developing efficient sampling techniques for overlapping populations. In such cases ingenious and workable methods are needed to take fullest advantage of the additional efficiency gained through making each successive sample contribute to the sample of the next concentric population.

The national cross section was made up of interviews allotted to approximately 450 localities in proportion to the population of geographic regions, states, and of size-of-town groups. Each interviewer was assigned quotas by sex, age, race, employment status and occupation. These quotas were based on 1940 Census data adjusted wherever possible to more recent Government estimates.

In four of the seven geographic regions, the national cross section sample was adequate. In the other three, additional interviews were taken to supplement those obtained in the national cross section. These extra interviews were assigned to the sample localities in proportion to the distribution of the general public in these regions.

Major industrial areas were defined as all counties in which a minimum of 3% of all employed men were engaged in the clients industry. These counties were stratified according to population and according to per cent of men employed in the industry. These substrata were set up so that the aggregate populations of the counties in each were approximately equal. A sample of counties was then drawn at random from each stratum. General-public interviews in the national or regional samples which were taken in counties where more than 3% of the employed men were engaged in this industry were included in the sample of major industrial areas. Consequently, the number of interviews allocated to counties in each stratum was reduced by the number of interviews already taken in this stratum in other samples. Since there was considerable overlap between the different areas analyzed, very substantial savings were effected by taking advantage of the duplications. Quotas were assigned by sex, age, race, employment status, and occupation in proportion to the distribution of such characteristics in each major industrial area.

Interviews with industry employees were taken in the sample counties in the major industrial areas in proportion to the employment in the industry in those areas. The remainder of the industry employee sample was allocated among the national sample localities in proportion to the employment in the industry in those localities.

Note: Quota sampling has been widely condemned in recent years. Principally, its validity can be questioned on the underlying inverse logic on which it is based—namely:

1. A representative cross section will reflect known characteristics accurately;
2. Therefore, a sample which reflects known characteristics accurately is a representative cross section.

In spite of this dubious logic, it can be shown that under certain not uncommon conditions quota sampling may have the highest cost efficiency of any method now known.

IX

A very interesting and ingenious application of an older technique, which has been largely abandoned in the past ten years, has recently been reported by Benson, Young, and Syze.[1]

This technique, now known as pinpoint sampling, was used by the American Institute of Public Opinion in connection with the 1944 presidential election.

Past election results for five particular states were found to represent a remarkable reflection of national election outcomes for all two-party contests since 1916.

Within each one of these five states, eight precincts were selected for the sample. The precincts within each state were chosen in such a way that, "giving equal weight to each, the average of the percentages of the vote of either major party was virtually the same as the election result for the state." This was found to have been true in three previous statewide elections.

Within the precincts selected interviewers were instructed to interview in every—or every other—dwelling unit, depending on the number of voters in the precincts. The interviewers still retained the freedom of selecting respondents within the family group, except for the limitation of obtaining over-all quotas for sex and age.

Interviews were completed in about one half of the households in the sample election districts. Slightly more than one fourth of the potential voters were interviewed.

The results show a deviation of 1.9% from the actual national vote. Part of this deviation was due to the selection of states and part to the selection of precincts. Benson, Young, and Syze concluded from the results that improvements could most profitably be made by enlarging the number of sample areas and taking a smaller number of interviews within each area.

---

[1] Edward G. Benson, Cyrus C. Young, and Clyde A. Syze, "Polling Lessons from the 1944 Election," *The Public Opinion Quarterly*, Winter 1945–46. The authors are indebted to Messrs. Benson, Young, and Syze for their permission to present their findings.

We have here a case of sampling by purposive selection.

> *Note:* The use of the method of purposive selection has been roundly condemned by many writers on the subject of sampling. Gini and Galvani in an article in *Annali Di Statistica* in 1929 showed empirically that this method was very inefficient. Later Neyman in his now famous article, "On the Two Different Aspects of the Representative Method," *The Journal of the Royal Statistical Society*, 1934, confirmed theoretically Gini's and Galvani's findings.
>
> Neyman, however, did express the opinion that under certain very rigid conditions purposive selection might be an efficient method of sampling. One of those very rigid conditions was that the technique might "work" when the controls were very highly and linearly correlated with the statistics to be measured. Although Benson, Young, and Syze do not discuss this point, it appears that this condition has been met. Therefore, it seems reasonable that the method of pinpoint sampling may have very important possibilities in predicting election results.

X

Although the problems already presented indicate the great variety of situations a commercial research agency has to meet, there are many more entirely different sampling and interviewing situations that cannot be covered in detail here.

Some of the sampling techniques described vary according to the particular survey situation. In a survey taken in a town and in its surrounding rural area, for neither of which block statistics were available, the sample was designed by drawing routes on a town map and state highway maps. The routes were drawn in such a way as to include different sections in the town as well as the different types of roads in the survey area. Interviewers were to interview in every *n*th house on the right side of the route designated until their quota was completed. Different starting points were chosen for the routes, some leading from the town towards outlying areas, others starting as far as 15 miles from the town and leading toward it.

Another example of a special situation: Even though interviews are generally done personally with only *one* respondent in any one household, there are cases where it is best to interview husband and wife together in order to approximate a family buying situation (for example, when a survey is made on intended purchases of refrigerators).

Even though most questioning is done by personal interview, there are situations where telephone interviews are satisfactory (for example, a survey with a short and strictly factual questionnaire and a specialized universe all members of which are telephone subscribers).

In other cases, such as in "in-plant studies," self-administered group interviews are used with a sample including every *n*th employee from the payroll.

In situations where respondents might feel that their statements could place them in some form of jeopardy, it may be necessary to interview on the street to allay any possible suspicion that respondents might be traced.

In cases of a management sample, such as sales managers or public-relations directors of companies with assets of more than $5,000,000, it may be necessary to sample from such sources as *Moody's* and *Standard and Poor*, and from individual state directories, supplemented by whatever information is available through local chambers of commerce.

In a survey with stockholders, where the composition of the universe was unknown, information was obtained on the characteristics of the types of people who are most often owners of stocks by including a question on stock ownership in other large national surveys. The survey with stockholders could then be restricted to these types, with the realization, of course, that the unusual types of stockholders would not be included.

There are numerous other situations—radio surveys, magazine surveys, advertising studies, and many others—which require various approaches suited to their special needs.

Looking back at the examples given, we realize that there are many other possible sampling approaches to these problems. The purpose of this paper is to report, not to analyze. These were the approaches that were actually used. They are presented here to illustrate a few of the many commercial uses of sampling.

VARIATION OF THE FREQUENCY OF FATAL QUARRELS WITH MAGNITUDE

LEWIS F. RICHARDSON
Hillside House, Kilmun, Argyll, Britain

A record of wars during the interval A.D. 1820 to 1945 has been collected from the whole world, and has been classified according to the number of war-dead. The smaller incidents have been the more frequent, according to a fairly regular graph which can be extended to quarrels that caused a single death.

INTRODUCTION

ANYONE WHO TRIES to make a list of "all the wars" (e.g. Wright [18] p. 636) encounters the difficulty that there are so many small incidents, that some rule has to be made to exclude them.

From the psychological point of view a war, a riot, and a murder, though differing in many important aspects, social, legal, and ethical, have at least this in common that they are all manifestations of the instinct of aggressiveness. This Freudian thesis has been developed by Glover [5], by Durbin & Bowlby [3], and by Harding [6]. There is thus a psychological justification for looking to see whether there is any statistical connection between war, riot, and murder. I dealt with this problem in 1941 by forming the inclusive concept of a "fatal quarrel," and then classifying the fatal quarrels according to the number of quarrel-dead [12]. By a *fatal quarrel* is meant any quarrel which caused death to humans. The term thus includes murders, banditries, mutinies, insurrections, and wars small and large; but it excludes accidents, and calamities such as earthquakes and tornadoes. Deaths by famine and disease are included if they were immediate results of the quarrel, but not otherwise. In puzzling cases the legal criterion of "malice aforethought" was taken as a guide.

The record of the number killed in a particular war is often uncertain by a factor of two. The meaningful part of the record can be separated from its uncertainty by taking the logarithm to the base ten and rounding the logarithm off to a whole number, or to the first decimal, according to the quality of the information. For simplicity I have lumped together the deaths on the opposing sides of the quarrel. The *magnitude* of a fatal quarrel is defined to be the logarithm to the base ten of the number of people who died because of that quarrel. The magnitude will be denoted by μ. The range of magnitude extends from 0 for a murder involving only one death, to 7.4 for the Second World War. Other well-known wars had magnitudes as follows: 1899–1902 British versus Boers 4.4; 1939–40 Russians versus Finns 4.83; 1861–65 North American Civil War 5.8. The magnitude of a war is usually known to within ± 0.2; so that a classification by unit ranges of magnitude is meaningful. These ranges have been marked off at 7.5, 6.5, 5.5, 4.5, 3.5, 2.5, 1.5 and perhaps at 0.5 and -0.5. Every fatal quarrel must lie inside one of these cells; it cannot lie on a boundary, because the antilogs of the boundary μ are not integers.

This abstract framework becomes of interest when the facts are sorted into it. As far as I have been able to ascertain, the historians have never sorted their facts by the scale of magnitude. The chief obstacle to any scientific study of wars-and-how-to-avoid-them, is each nation's habit of blaming other nations. National prejudice can however be avoided by taking the whole world as the field of study; and by taking a time-interval longer than personal memory. For this reason I have made a search for the records of fatal quarrels in the whole world since the beginning of A.D. 1820. The task has been long; and the results are here presented in brief summary. For magnitudes greater than 2.5, the facts were mostly obtained from works on history. For magnitudes less than 0.5 they were taken from criminal statistics. For magnitudes between 0.5 and 2.5 the information is scrappy and unorganized; what there is of it suggests that such small fatal quarrels were too numerous and too insignificant to be systematically recorded as history, and yet too large and too political to be recorded as crime.

It is important to know whether the facts have all been gathered. The best evidence is provided by the progress of search. I began in the year 1940. At first my collection grew rapidly, then slowly; then various revisions caused its totals to oscillate slightly. The collection appeared to be sufficiently complete to warrant the publication of a summary, which appeared in *Nature* of 15th November, 1941. Afterwards the publication of Quincy Wright's list of wars ([18] Appendix XX) provided a stimulus to further enquiry, which involved the consultation of some seventy history-books. In the following table three stages of my records are compared. All refer to fatal quarrels which ended from A.D. 1820 to 1929 inclusive.

Date and place of record	Ends of range of magnitude			
	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$
	Numbers of fatal quarrels recorded			
1941. *Nature* of Nov. 15	1	3	16	62
1944. J.R.S.S. 107, 248	—	—	—	63
1947. Here.	1	3	20	60

It is seen that there has been a sort of convergence, which makes the present numbers worthy to be discussed. That is for $\mu > 3.5$. For smaller incidents in the range $2.5 < \mu < 3.5$ my collection has continued to grow; the sign $\geqq$ is therefore prefixed to its totals.

The world-total for the murders was obtained from the murder-rates of 17 countries. They were weighted by populations. The weighted mean rate for the world was found to be 32 murders per million per year. This is not as definite as the statistics of wars, because the criminal statistics of China, Africa, and South America (except Chile) were not available. But even if that estimate were as much as three times wrong, which is incredible, it still would give significant results in Figures 0 and 1, because they relate to such enormous ranges. The evidence which I had concerning time-changes in the murder-rate was scrappy and conflicting; so that provisionally I took the rate to be constant. The mean population of the world over the 126 years from A.D. 1820 to 1945 was computed from a publication by Carr-Saunders [1A] and was found to be 1.49×10^9. So the world-total of murders was computed as $1.49 \times 10^9 \times 32 \times 10^{-6} \times 126 = 6.0 \times 10^6$.

A revision based on the murder-rates of 21 countries, gave a mean rate of 37 instead of 32, and a world-total of 7 instead of 6 million. The following diagrams and tables were based on the earlier estimate and have not been altered.

A CONSPECTUS WHICH SUITS THE WARS

The number of fatal quarrels has been counted, or estimated, in unit ranges of magnitude, with the following results for the world as a whole. The date of a quarrel is taken to be that of the termination of hostilities. The interval is A.D. 1820 to 1945 inclusive.

Ends of range of magnitude	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$	$3\pm\frac{1}{2}$ · · · Murders
Observed number of fatal quarrels	2	5	24	63	$\geqq 188$ · · · 6×10^6

It is desirable to show all these facts together on a single diagram. If the number of fatal quarrels per unit range of magnitude were taken as the ordinate, then either the wars would show, or the murders, but not both. When however the logarithm of the number of fatal quarrels per unit range of magnitude is taken as the ordinate, then all parts show equally well. Each murder presumably involved a small number of deaths, such as 1, 2, 3, or 4, so that the corresponding logarithms were 0.000, 0.301, 0.477, or 0.602. How to group them was not obvious. As a first expedient I extended to the murders the system of classification which had been found to suit the wars, namely by unit ranges of magnitude cut at 7.5, 6.5, 5.5, 4.5, 3.5, and so on by extension to 0.5 and -0.5. Murders involving 4 or more deaths were regarded as negligibly rare. The ordinate f is defined to be

$$f = \log_{10}\left\{\frac{\text{number of fatal quarrels}}{\text{corresponding range of magnitude}}\right\}.$$

This rough and ready scheme gives the conspectus shown in Figure 0. It is the type of diagram with which the author began, and it has now been asked for by a referee. A dotted curve has been drawn across the gap to suggest that there is statistical continuity between the wars and the murders.

THE FAULTS OF FIGURE 0 AND HOW TO AMEND THEM

In order to match the wars, the murders have been enclosed in a unit range of magnitude, as shown at A in Figure 0. But the half of this range, from $\mu = -0.5$ to μ just less than zero, is necessarily empty. Would it be more reasonable to regard the murders as enclosed in a half-unit range? If so the line representing them in Figure 0 would be halved in length. Also the number of murders per unit of magnitude would be doubled. So the representative line would be raised by $\log_{10} 2$ into position marked B. Worse ambiguities occur if we wish to show separately the number of murders each causing one death. Let there be M of them. To what range of magnitude ought they to be attributed? The range must contain the point $\mu = 0$, and must not contain the point $\mu = \log_{10} 2$, but otherwise is unspecified. Yet M has to be divided by this indefinite range in order that the reckoning may be "per unit range of magnitude". Figure 0, though suitable for the wars, is too vague for quarrels that caused only one death.

It may be thought that the root of the trouble is the reckoning "per unit range of" anything, and that if ranges were abandoned then murders and wars could be compared by plotting the logarithm of their numbers against their magnitudes. Such a diagram would consist of

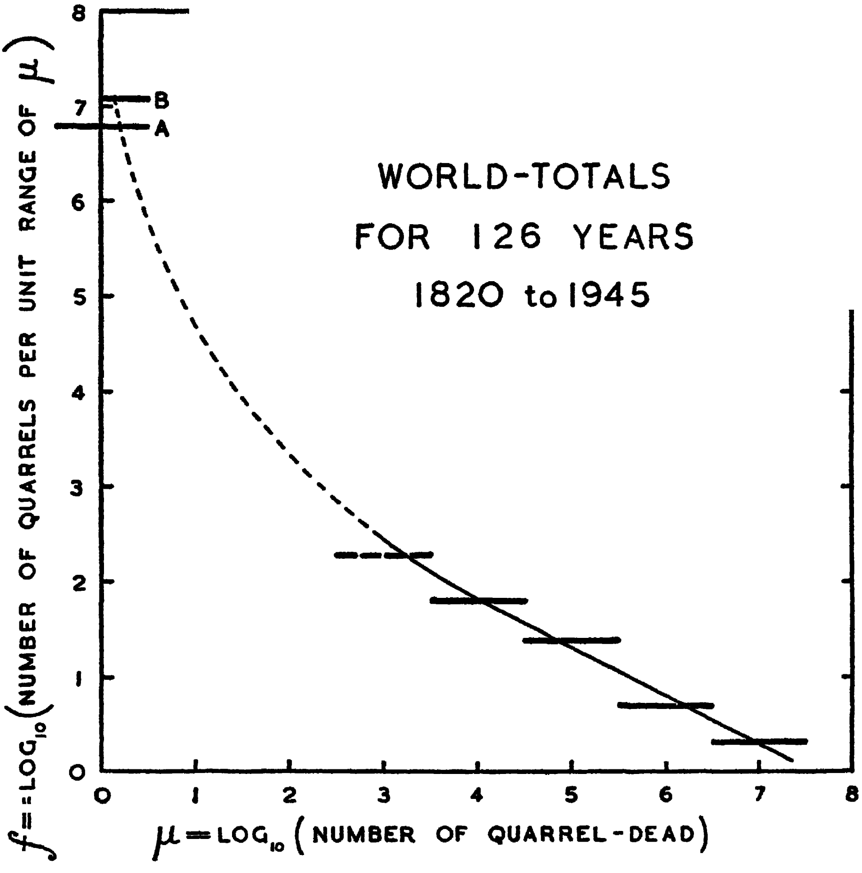


FIGURE 0

points, not segments. There are two objections to this proposal: (i) the magnitudes of wars are imperfectly known and some of them are rounded off to whole numbers; so that there would be many artificially coincident points, (ii) if the magnitudes were better known they would presumably be scattered, so that a point-diagram of wars would present great irregularities, of no interest. For when n exceeds 1000 there is never likely to be any statistically significant distinction between wars involving n and $n+1$ deaths. In brief we must have points for the murders and averages over ranges for the wars, and yet these diverse representations must somehow be compared. The solving idea, which melted all the obstacles, was that the appropriate measure of any range of quarrel-dead is simply the number of integers in that range, and not the corresponding range of magnitude, nor any other artificial conception.

A DIAGRAM WHICH SUITS THE MURDERS

The logically simple idea is that there were, say $q(n)$ quarrels each involving n deaths. (1)

For example in England and Wales during 1935 and 1936 there were known to the police 186 cases of murder of 213 persons aged over one year. In 74 cases (involving 91 persons) the murderer or suspect committed suicide. In 105 cases (115 victims) 111 persons were arrested. In the remaining 7 cases, involving 7 victims, no arrest was made. Of the 111 persons arrested 16 were executed (Whitaker's Almanacks 1938 and 1939, p. 652). These statements can be interpreted, with only slight uncertainty, by the following distribution.

				totals
$n=$	1	2	3	
$q(n)=$	92	71	23	186
$nq(n)=$	92	142	69	303

The number of deaths, n, connected with the quarrel is here taken to include, along with those of the victims, also those of the murderer or suspect, whether by suicide or execution. It seems permissible, for illustration until better information can be found, to regard the 6 million murders in the world from 1820 to 1945 as the number of fatal quarrels for $n=1$, 2 and 3, and to subdivide this total among its three parts in the same ratios as the total for England and Wales is subdivided. Thus one finds for the world, the following enlarged values of $q(n)$. To take one country as a sample of the world is very unsatisfactory. But it should be noted that the uncertainties of these $q(n)$ are insignificant in comparison with the vast range of Figure 1 where $q(n)$ varies in a ratio exceeding 10^{13}.

				totals	
$n=$	1	2	3		
$q(n)=$	3.0	2.3	0.7	6.0	millions
$nq(n)=$	3.0	4.6	2.1	9.7	

These give three points in the top left corner of Fig. 1, with the coordinates

$\log_{10} n=$	0	0.301	0.477
$\log_{10} q(n)=$	6.48	6.36	5.85

A GENERAL AND WELL DEFINED CONSPECTUS

This will now be obtained by extending to the wars the method introduced above for the murders.

In the region of wars, say for $n > 1000$, the function $q(n)$ is usually zero and occasionally unity; so that a diagram of it would be bristly and unreadable. Therefore for the wars we should take an average of $q(n)$, say $\bar{q}(n)$. As a preliminary to the definition of $\bar{q}(n)$, it must be understood that the fatal quarrels are arranged in order of n. The definition of $\bar{q}(n)$ is then

$$\bar{q}(n) = \frac{\text{number of fatal quarrels in a range of } n}{\text{number of integers in that range of } n}. \tag{2}$$

Briefly, and somewhat inaccurately, $\bar{q}(n)$ may be called the "average frequency of quarrels per unit range of quarrel-dead." As in all frequency-diagrams so here the choice of the group-range has to be a compromise: if the range is too small, the diagram is spiky; if the range is too large, essential features are flattened out. The question of the best range must be decided by trial.

The diagram of q and $\bar{q}$ as functions of n is the logically simple conspectus of the frequency of wars and murders. This diagram cannot be drawn, because no sheet of paper is large enough to show the facts at both ends of the range. The difficulty is overcome by plotting the logarithms of q, $\bar{q}$ and n. The magnitude μ has already been defined by

$$\mu = \log_{10} n. \tag{3}$$

Another oft-recurring symbol, ϕ, must now be defined.

Let $\log_{10} \bar{q}(n) = \phi(\mu)$, say, for the wars: while for the murders

$$\phi(\mu) = \log_{10} q(n) \text{ without averaging.} \tag{4}$$

The last column of the following table shows $\phi(\mu)$, apart from some corrections for grouping, which will be explained on p. 542. The first and last columns of this table provide the abscissa and ordinate of the accompanying diagram (Figure 1), as far as the wars are concerned. It is evident that the wars can be satisfactorily fitted by a straight sloping line. On account of grouping, the wars are represented, not by points

Range of magnitude μ	Number of integers in range of war-dead	Observed number of fatal quarrels Years 1820–1945	$\log_{10}$ (number of quarrels per unit range of war-dead)
2.5 to 3.5	2,846	$\geqq 188$	$\geqq -1.180$
3.5 to 4.5	28,460	63	−2.655
4.5 to 5.5	284,605	24	−4.074
5.5 to 6.5	2,846,050	5	−5.755
6.5 to 7.5	28,460,499	2	−7.153

as the murders are, but by horizontal segments. It will be shown later (p. 542) that the $\phi(\mu)$ graph should pass slightly above the mid-points of these segments. That is to say the sloping straight line was actually obtained by a second approximation.

The relation between the ordinates of Figures 0 and 1

The abscissa μ is the same in both diagrams and the ordinates are respectively f and ϕ, defined thus

$$f = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{\log_{10} n_b - \log_{10} n_a}\right\}$$

where $n_1 - 1 < n_a < n_1$, and $n_1 + s < n_b < n_1 + s + 1$. These extra bits at the ends of the range have to be put in to prevent f becoming infinite at $s=0$.

$$\phi = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{s+1}\right\}$$

which is definite at $s=0$. By subtraction

$$f - \phi = \log_{10}\left\{\frac{s+1}{\log_{10} n_b - \log_{10} n_a}\right\}.$$

This expression involves an ambiguity, because n_b is loose in a range of unity, and so is n_a. The practical question however is whether the ambiguity is greater than the uncertainty of the observations. It is necessary to distinguish different cases.

Case (i). The range here used in treating the wars is $\mu_0 - \frac{1}{2} \leqq \mu \leqq \mu_0 + \frac{1}{2}$ where $\mu = \log_{10} n$ and where $\mu_0 \geqq 3$. In these circumstances arithmetic shows that

$$f - \phi = \mu_0 + 0.454, \tag{5}$$

reliable to the third decimal. Seeing that the μ of a particular war is often uncertain by ±0.1, equation (5) is admirably accurate for the purpose.

Case (ii). We may conceivably wish to consider a shorter range $\mu_0 - \epsilon \leqq \mu \leqq \mu_0 + \epsilon$ where μ_0 is still in the region of the wars. If ϵ is small, but not too small, there is a quasi-limit such that

$$f - \phi = \mu_0 + 0.362. \tag{5A}$$

This was obtained by approximating to the difference-ratio by a derivative, and was confirmed by arithmetic.

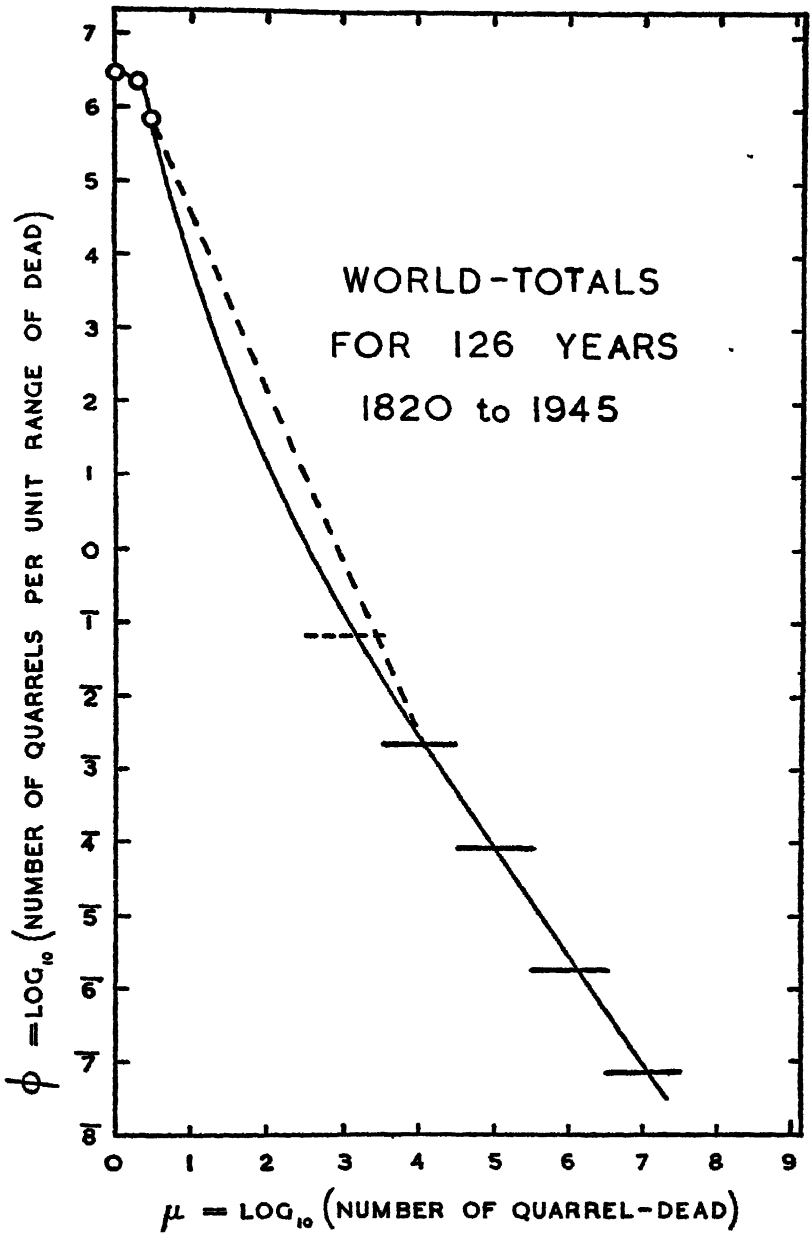


FIGURE 1. THE WHOLE RANGE OF FATAL QUARRELS

THE HORIZONTAL SCALE EXTENDS FROM A SINGLE PERSON ON THE LEFT TO THE POPULATION OF THE WORLD ON THE RIGHT. THE THREE DOTS AT THE LEFT HAND TOP REPRESENT THE MURDERS. THE SHORT HORIZONTAL LINES REPRESENT THE WARS GROUPED BY MAGNITUDES. ONE OF THESE LINES IS DOTTED, BECAUSE IT SHOULD BE RAISED TO AN UNKNOWN HEIGHT.

Case (iii). If there are only a few integers n in the range then the ambiguity becomes troublesome. This is the situation as to the murders. The ambiguity of $f-\phi$ is more than that of some criminal statistics. The fault lies in f, not in ϕ. When $n_1=1$ and $s=0$ it is hardly possible to locate n_a and n_b by any reasonable convention, as already explained on page 526. To revert to Cases (i) and (ii), either of them give the same relation between the slopes of the f- and ϕ-diagrams. As these slopes are in fact negative, it is convenient to write the relation thus

$$-\,d\phi/d\mu_0 = -\,df/d\mu_0 + 1 \tag{5B}$$

which shows that the ϕ-diagram is the steeper.

The slope in the region of the wars

This was found, by the principle of Maximum Likelihood, [4 & 10] to be

$$d\phi/d\mu_0 = -\,1.50. \tag{5C}$$

Interpolation across the gap between the wars and the murders

The dotted horizontal segment for the small wars represents only a lower bound, and otherwise should be ignored. Between the definite end-points of the gap, the mean slope is

$$d\phi/d\mu = -\,2.38. \tag{6}$$

This is shown by a straight line of dashes.

As to the choice of a curve, there are the general clues of simplicity and of continuous slope. For simplicity a circular arc was chosen; and, for continuous slope, the arc was made tangent to the straight line which fits the wars. It was then found, as an independent confirmation that the slope of the circular arc joins on continuously to that for the murders. Some collateral evidence is provided by banditry.

A SEARCH FOR THE FACTS ABOUT QUARRELS THAT CAUSED FROM 4 TO 315 DEATHS $(0.5<\mu<2.5)$

That there were many small fatal quarrels is shown by a statement made by Quincy Wright [18] in the introduction to his long list of "Wars of Modern Civilization." He remarked that "A list of all revolutions, insurrections, interventions, punitive expeditions, pacifications, and explorations involving the use of armed force would probably be more than ten times as long as the present list, . . . " Wright however did not classify fatal quarrels according to the number of quarrel-dead. Similarly whilst reading histories and police-reports and listening to radio-news I have noticed allusions to very many incidents less important than wars, but more important than murders. References [2], [7], [11], [15], contain some remarkable examples. However I see no hope of obtaining world-totals of the numbers of such incidents during the 126 years to which Figure 1 relates, and classified moreover by magnitudes in the range $0.5 < \mu < 2.5$. Any factual test of the interpolated portion of Figure 1 will have to depend on smaller samples. The record of banditry in Manchoukuo is comparable on certain assumptions, and that of gangsters in Chicago on further assumptions. These ordered collections of facts, being rare specimens, are correspondingly valuable.

BANDITRY IN MANCHOUKUO DURING THE YEAR 1935

This is of especial interest, because the facts are given in a form which throws light on aggregation for aggression. The following is quoted from the *Japan and Manchoukuo Year Book*, 1938, pp. 692–95.

"At the time of the founding of Manchoukuo in March of 1932 the total number of bandits exceeded 100,000. By September of the same year the number had increased to 210,000 due principally to the subversive activities of Chang Hsueh-Liang's remnant troops who were thrown out of employment following the downfall of the young marshal. Since then, however the number of such bandits has been on the decrease as a result of their suppression by Manchoukuo and Japanese forces.

"Compared with the condition obtaining in 1932 two factors loom in prominence with regard to the bandit situation. Firstly may be noted the actual reduction of bandits as a whole, and, secondly, the shrinkage in size of bandit groups. In 1932 some bandit groups had an actual fighting force of 30,000 men, but at present the average is below 50 bandits per group. The chief cause for the existence of bandits in Manchoukuo is believed to be an economic one, resulting from unemployment."

STATISTICS OF BANDITRY

Size of bandit groups						
1 to 30	31 to 50	51 to 100	101 to 200	201 to 300	301 to 500	>501
Corresponding number of raids						
28,145	4,784	3,864	1,530	455	240	130

The first step towards making a comparison with larger fatal quarrels is to compensate for the unequal ranges of size in the above quotation, by dividing each number of raids by the corresponding range of size. For example $4784/(50-31+1)=239.2$. Any such ratio may be called the "number of raids per unit range of membership." For the first six groups it runs as follows

938.17 239.2 77.28 15.30 4.55 1.20

and shows plainly that *the smaller incidents were much the more frequent.*

Was a raid like a small battle or like a small war? In other words did the same group of bandits perpetrate many raids? This can be answered by attending to totals. The total number of appearances of bandits can be underestimated from the above table by multiplying the least size in each column by the corresponding number of raids, and comes to 758,371 or more. A moderate estimate is 1.3 million.

But the total number of bandits is stated to have been much less than 210,000. It is evident therefore that on the average the same bandit appeared in several raids. That is to say *a raid resembled a battle rather than a war.*

In the same year and same region the numbers killed in connection with banditry are stated thus: bandits 13,338, suppression troops 1361, civilians 2512. The total dead was therefore 17,211. This is 1.3 per cent of the total number of appearances of bandits.

The facts about banditry are plotted in Figure 2. This is not a $\phi(\mu)$ diagram, yet it is rather like one. For the abscissa, ν say, is $\log_{10}$ of the number of bandits in the group, whereas μ is $\log_{10}$ of the number of quarrel-dead. Again the ordinate, ψ say, is $\log_{10}$ of the average number of raids per unit range of membership, whereas ϕ is $\log_{10}$ of the average number of fatal quarrels per unit range of quarrel-dead. The first group in the data has a membership range of from 1 to 30, which is in a ratio too great for present purposes. Although the first group is shown on Figure 2, it will be ignored in the discussion. The blunted top of Figure 1 has however some resemblance to the horizontal in Figure 2. The other groups are well fitted by a straight line having the slope

$$d\psi(\nu)/d\nu = -2.29. \tag{7}$$

This has a remarkable resemblance to the mean slope, −2.38, across the gap in the world-diagram, as stated in (6). Can this agreement of the slopes be a mere coincidence? Or is it a clue to a general law concerning aggregation for aggression? It is at any rate easily explained by the following two assumptions:—

Firstly let us suppose that the quarrel-dead were on the average proportional to the number of bandits in the group. That is to say that groups continued to make raids until a constant fraction of their members had been killed. The supposition is that

$$n = A \times \text{(number of bandits in group)}$$

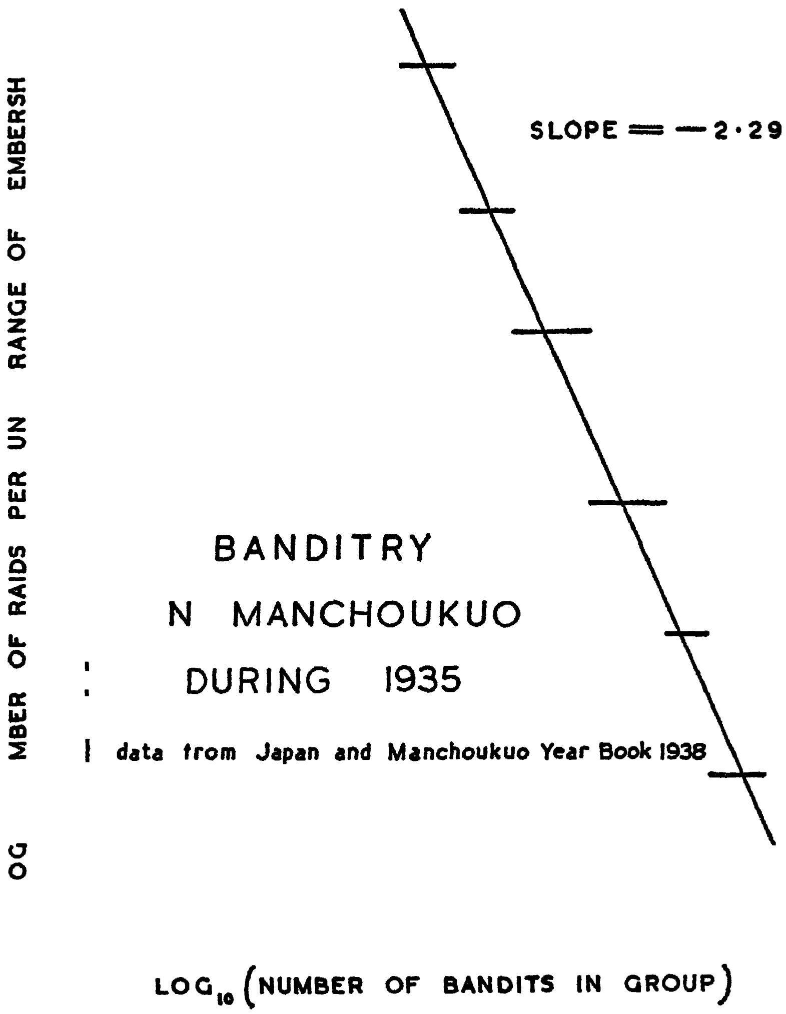


FIGURE 2

where A is independent of the size of the group. Accordingly

$$\mu = \nu + \log_{10} A. \tag{8}$$

It may not be necessary thus to suppose that the percentage of killed was exactly the same for every group, but only to suppose that it had no correlation with the size of the group.

Secondly let us suppose that the number of raids which a group of bandits made, as part of their quarrel with the rest of the community, was not correlated with the size of the group. More precisely the supposition is that:

$$(\text{number of raids made by a group}) = Bq(n)$$

where B is independent of n, so that

$$\psi(\nu) = \phi(\mu) + \log_{10} B. \tag{9}$$

According to these assumptions the $\psi(\nu)$ graph, when slid parallel to itself, horizontally through log A and vertically through log B, becomes a *local* $\phi(\mu)$ graph. The remarkable agreement of the slopes can therefore be interpreted as meaning that one of the characteristics of aggregation for aggression was the same in Manchoukuo as it was in the world-total.

That broad agreement is interesting; but there remain some doubts or discrepancies in detail. The Manchurian data do not fix A and B, though A can be roughly estimated. For it has been shown that on the average 1.3 per cent of the bandits who appeared in a raid were killed, and that a group on the average perpetrated several raids. The total quarrel-dead, n, may have been about 10 per cent of the number of bandits in the group; that is to say $A=0.1$, so that $\mu=\nu-1$. The mean value of μ on the straight part of the bandit-diagram is accordingly about one. If so, there is a discrepancy with the circular arc on the world-diagram. For at $\mu=1$ the slope of the arc is decidedly steeper, namely $d\phi/d\mu=-3.1$. The Manchurian data therefore suggest that the gap in the world-diagram (Figure 1) should be closed, not by the circular arc, but by the straight segment of slope -2.38 leaving discontinuities of slope at its two ends. I regard that as a suggestion to be remembered, but not to be acted on without further evidence. The shift $\mu=\nu-1$ cuts off most of the horizontal part of Figure 2 and thus increases its resemblance to the top of Figure 1.

GANGING IN CHICAGO

F. M. Thrasher [16] made a study of 1313 gangs in Chicago with a view to their re-direction, or to some other social improvement.

He defined a gang in these words (p. 57): "The gang is an interstitial group originally formed spontaneously, and then integrated through conflict. It is characterized by the following types of behaviour: meeting face to face, milling, movement through space as a unit, conflict, and planning. The result of this collective behavior is the development of tradition, unreflective internal structure, *esprit de corps*, solidarity, morale, group awareness, and attachment to a local territory."

Thrasher gives (p. 319) a table showing the approximate numbers of members in 895 gangs. The first two columns of the following table are copied from Thrasher, the third column is a deduction, designed to smooth out the disparities of the given ranges. The number of members in a gang is the same sort of quantity as the number of bandits in a group; so the same symbol ν is here used for $\log_{10}$ of the first column. Log_{10} of the third column is here denoted by $\chi(\nu)$ to distinguish it from $\psi(\nu)$, because the number of gangs is not similar to the number of raids, though related to it. Figure 3 shows a graph of $\chi(\nu)$. Apart from gangs of 15 or less, which show again the blunted top, the $\chi(\nu)$ graph is well fitted by the straight line shown, which has a slope

Number of members in gang	Number of such gangs	Gangs per unit range of membership
inclusively		
3 to 5	37	12.3
6 to 10	198	39.6
11 to 15	191	38.2
16 to 20	149	29.8
21 to 25	79	15.8
26 to 30	46	9.2
31 to 40	55	5.5
41 to 50	51	5.1
51 to 75	26	1.04
76 to 100	25	1.00
101 to 200	25	0.25
201 to 500	11	0.37
501 to 2,000	2	0.0013

$$d\chi(\nu)/d\nu = -2.30. \tag{10}$$

It is remarkable that this slope is almost the same as $d\psi(\nu)/d\nu = -2.29$ for banditry in Manchoukuo. This agreement strengthens the suspicion that some fairly general tendency concerning aggregation for aggression is revealed by these otherwise scattered phenomena. Hypothetical explanation is again easy. The distinction between the number of gangs and the number of raids would not affect the slope, if it were true that the number, C say, of raids made by a gang was independent of the size of the gang. For then

$$\psi(\nu) = \chi(\nu) + \log_{10} C \tag{11}$$

and the constant $\log_{10} C$ would disappear on taking $d/d\nu$. Yet several features remain unclear. The Chicago gangs were certainly aggressive.

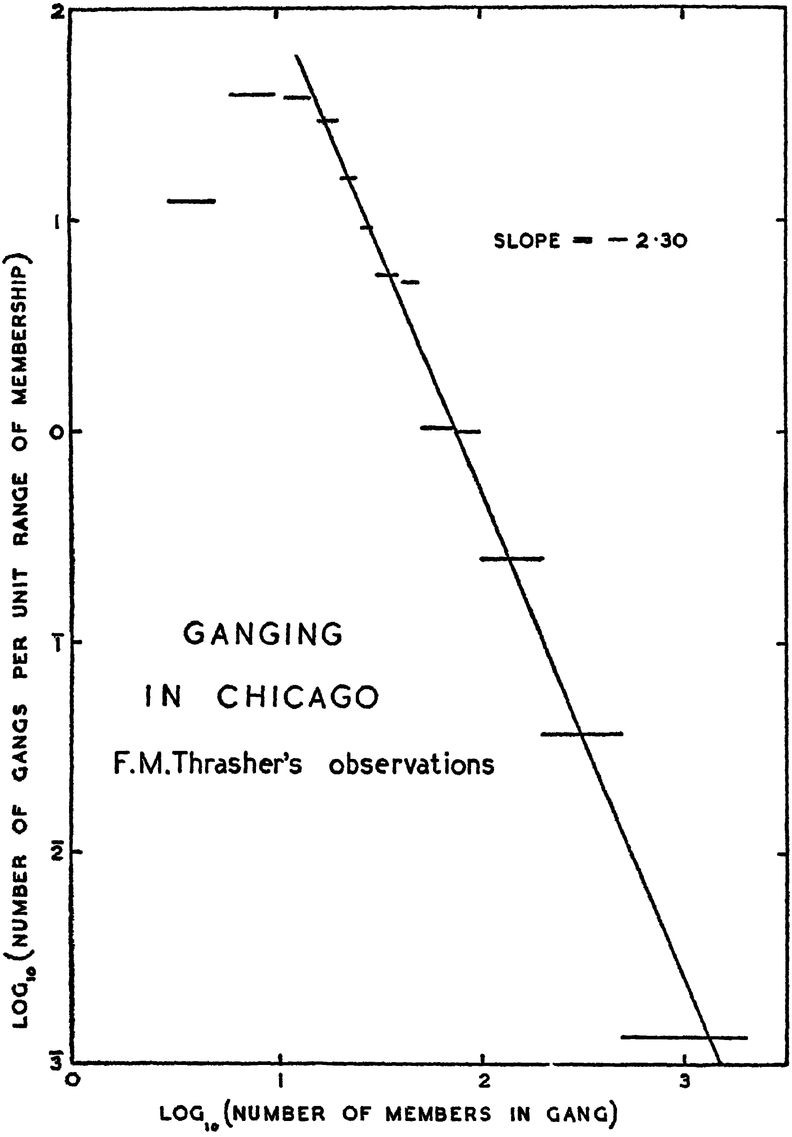


FIGURE 3

Thrasher has a chapter headed "Gang Warfare" and beginning with these words

> "The gang is a conflict group. It develops through strife and thrives on war-

> fare. The members of a gang will fight each other. They will even fight for a 'cause.' . . . Gangsters are impelled, in a way, to fight; so much of their activity is outside the law that fighting is the only means of avenging injuries and maintaining the code."

Thrasher mentioned a few gang-fights which went as far as homicide, but he did not give any comprehensive statistics of the quarrel-dead. I am indebted to Prof. Lundberg and Margaret Black Richardson, for the indication of Thrasher's work.

SUMMARY

The $\phi(\mu)$ diagram in Figure 1 is the best conspectus of the facts for the world as a whole. Alternative tracks are shown across the gap where world-totals are lacking. Samples from Manchoukuo and Chicago confirm the slope of the straight alternative, whereas continuity of slope is a consideration in favor of the circular arc.

MORE EXACT CONNECTION BETWEEN THE $\phi(\mu)$ GRAPH OF FIGURE 1 AND THE NUMBERS OF FATAL QUARRELS IN UNIT RANGES OF MAGNITUDE. INTERPOLATION ACROSS THE GAP.

We need to connect the mean over a range with the value at its midpoint. It might be thought that the connection was obvious; for a glance at Figure 1 shows that the graph is fairly straight; and of course the mean ordinate of a straight segment is simply the ordinate at its midpoint. But that would be a delusion based on a doubly wrong type of mean. For here the mean of q, not of the ordinate $\log_{10} q$, has to be taken over the integers in the range of n, and not over the abscissa $\mu = \log_{10} n$. It will be shown that the logarithmic transformation of both coordinates introduces correcting factors which depend on the slope, vanishing when it is either 0 or -2, but otherwise often considerable. They are obtained from the simplest formula that can fit the facts tolerably, namely the formula for a straight line on the $\phi(\mu)$ graph. The straight lines are suitably short segments; for it would be a mistake to smooth away significant detail by the over-wide sweep of any too-simple formula. In particular the sharp bend near the top of Figure 1 comes from British criminal statistics where it appears to be significant; and something rather like it is to be seen in the graphs which represent banditry in Manchoukuo and ganging in Chicago; so this bend should certainly not be smoothed away. But elsewhere a short length of the $\phi(\mu)$ graph can be represented by

$$\phi(\mu) = B - C\mu$$

where B and C are positive constants. On taking antilogarithms this becomes

$$q = 10^B n^{-c}$$

an inverse power law. To find $\bar{q}$ this n^{-c} has to be averaged over a range of n. As n proceeds only by integers, the average should be made by sums, not by integrals. However the integral $\int n^{-c}dn$ gives a useful approximation to the sum.

This programme will now be worked out in proper detail.

The notation is that already defined in connection with the graph. The typical range ends, say, at $\mu = \mu_0 \pm \frac{1}{2}$, where $\mu_0 = 7, 6, 5, 4, 3, 2, 1$. It is convenient to begin with a rough and ready approximation, R, to the number Q of fatal quarrels in the range, whereby $\bar{q}(n)$ at the midpoint is multiplied by the number of integers in the range of quarrel-dead, which it is convenient for a later purpose, to denote by $\sum_n 1$ where $\sum_n$ sums for all n in the range of μ. Accordingly

$$R = 10^{\phi(\mu_0)} \sum_n 1. \tag{12}$$

For example, as read from the circular arc on Figure 1:

Range of magnitude $\mu = \log_{10} n$	Ordinate at midpoint $\phi(\mu_0)$	$\therefore \bar{q}(n)$ at midpoint	Number of integers in the range of quarrel-dead $\sum_n 1$	First approximation to number of quarrels R
0.5 to 1.5	3.87	7410	28	207,000
1.5 to 2.5	1.23	17.0	285	4,850
2.5 to 3.5	$\bar{1}.17$	0.131	2846	373

The above easy method is however suspect, because $\bar{q}(n)$ varies in the ratio 1740 in one of the ranges thus represented by its midpoint. So we must consider corrections.

Correction for width of range

As the $\phi(\mu)$ graph is nearly or quite straight, let us represent a local portion of it by

$$\phi(\mu) = \phi(\mu_0) - c(\mu - \mu_0) \tag{13}$$

where μ_0 is the midpoint of the range and $c = -d\phi/d\mu$ there. The distinction between $q(n)$ and $\bar{q}(n)$ is one that concerns detailed observations. It does not arise when, as now, deductions from a smooth curve in question. From (13)

$$q(n) = 10^{\phi(\mu)} = 10^{\phi(\mu_0)+c\mu_0}n^{-c}. \tag{14}$$

Let Q denote the total number of fatal quarrels in the range, then

$$Q = \sum_n q(n) = 10^{\phi(\mu_0)+c\mu_0}\sum_n n^{-c} \tag{15}$$

So from (12) and (15)

$$Q = R10^{c\mu_0}\sum_n n^{-c} \Big/ \sum_n 1. \tag{16}$$

The coefficient of R will be called the correcting factor. Two verifications of (16) may be noted: (i) if $c=0$, then $Q=R$. (ii) if the range of quarrel-dead contained only a single integer n, and if it were located at μ_0 so that $10^{\mu_0}=n$, then (16) would become $Q=Rn^c n^{-c}/1=R$, as it should.

The sum in (16) was evaluated by a method connected with the "integral test" for the convergence of series, in the following manner. Let

$$b = 10^{\mu_0+1/2} \quad \text{and} \quad a = 10^{\mu_0-1/2} \text{ for brevity.} \tag{17}$$

In the following chain of equations, the first is an approximation which improves as μ_0 increases, and the other are accurate.

$$\sum_n n^{-c} \simeq \int_a^b t^{-c}dt = \frac{1}{1-c}(b^{1-c} - a^{1-c})$$

$$\begin{matrix}10^{\mu_0(1-c)}\\ 1 -\end{matrix}\left\{10^{(1-c)/2} - 10^{(c-1)/2}\right\}. \tag{18}$$

In the special case of $c=0$, (18) gives the approximation

$$\sum_n 1 = 10^{\mu_0}(10^{1/2} - 10^{-1/2}) = b - a. \tag{19}$$

Strictly $\sum_n 1$ is the integer next below $b-a$; but it is advisable to use the same type of approximation in both numerator and denominator of (16). Insertion of (18) and (19) into (16) gives the correcting factor

$$Q/R = \frac{1}{1-c}\,\frac{10^{(1-c)/2} - 10^{(c-1)/2}}{10^{1/2} - 10^{-1/2}}. \tag{20}$$

This expression is unity when $c=0$ or $c=2$.

For the wars, $c=1.50$ and $Q/R=0.855$.

The observed quantity is Q in the unit range of magnitude. The $\phi(\mu)$ curve should however be connected to μ at a point, and so to R. We have

$$\log_{10} R = \log_{10} Q - \log_{10} 0.855 = \log_{10} Q + 0.068.$$

In the $\phi(\mu)$ graph (Figure 1) the sloping line for the wars was therefore raised 0.068 above that which best fits the observations grouped in unit ranges of magnitude.

In the unexplored region we have, on taking the first approximation R from the previous table:

World-totals for the 126 years 1820 to 1945

Range of magnitude μ	$c = -\frac{d\phi}{d\mu}$	Correcting factor Q/R	Corrected number of fatal quarrels Q	Range of number killed in a quarrel n
0.5 to 1.5	3.14	1.92	397,000	4 to 31
1.5 to 2.5	2.32	1.16	5,630	32 to 316
2.5 to 3.5	1.87	0.95	354	317 to 3162

The fourth column shows the improved estimates for the unexplored region between the wars and the murders according to the circular arc in Figure (1). The difficulties of direct counting have already been described, and they are emphasized by these large numbers. In the range where 354 are expected, I have counted 188. If the straight line in Figure (1) were accepted instead of the circular arc, 354 would be increased to 2530. Although I know that the search is incomplete, I am unable to believe that less than a tenth have been found; and so I prefer the circular arc.

THE TOTAL NUMBER OF PERSONS WHO DIED BECAUSE OF QUARRELS DURING THE 126 YEARS FROM 1820 TO 1945 A.D.

Ends of range of magnitude	Total number of deaths in millions	How computed
$7\pm\frac{1}{2}$	36	By summation over a list of fatal quarrels.
$6\pm\frac{1}{2}$	6.7	
$5\pm\frac{1}{2}$	3.4	
$4\pm\frac{1}{2}$	0.75	
$3\pm\frac{1}{2}$	0.30	From the circular arc in Figure (1) by the method described below.
$2\pm\frac{1}{2}$	0.40	
$1\pm\frac{1}{2}$	2.2	
$0\pm\frac{1}{2}$	9.7	From page 528
	Total 59	

A remarkable feature of the above table is that the heavy loss of life occurred at the two ends of the sequence of magnitudes, namely the World Wars and the murders. The small wars contributed much less to the total. The total deaths because of quarrels should be compared with the *total deaths from all causes.* There are particulars given by de Jastrzebski [9], and in the Statistical Year Books of the League of Nations, which allow the total to be estimated. A mean world population of 1.5×10^9 and a mean death rate of 20 per thousand per year would give during 126 years 3.8×10^9 deaths from all causes. Of these the part caused by quarrels was 1.6 per cent. This is less than one might have guessed from the large amount of attention which quarrels attract. Those who enjoy wars can excuse their taste by saying that wars after all are much less deadly than disease.

The method of computation in the unexplored region was an extension of that explained above. The number D of dead in any range of μ is strictly

$$D = \sum_n nq(n) = \sum_n 10^{\mu+\phi(\mu)}. \tag{21}$$

First a common-sense estimate, E of D was obtained from the graph at the midpoint μ_0 of the range $\mu_0-\frac{1}{2}\leqq\mu\leqq\mu_0+\frac{1}{2}$ thus, by (12),

$$E = 10^{\mu_0}R = 10^{\mu_0+\phi(\mu_0)}\sum_n 1. \tag{22}$$

Then E was corrected by multiplication by a factor which was found to be approximately

$$\frac{D}{E} = \frac{1}{2-c}\left\{\frac{10^{(2-c)/2} - 10^{(c-2)/2}}{10^{1/2} - 10^{-1/2}}\right\}. \tag{23}$$

The theory of this correction is based on the approximation (13), whence it follows that

$$\mu + \phi(\mu) = c\mu_0 + \phi_0 + \mu(1-c). \tag{24}$$

From (21) and (24)

$$D = 10^{c\mu_0+\phi(\mu_0)}\sum_n n^{1-c}, \quad \text{and} \tag{25}$$

$$\frac{D}{E} = 10^{\mu_0(c-1)}\,\frac{\sum_n n^{1-c}}{\sum_n 1}. \tag{26}$$

The sums in (26) were obtained from (18) by suitable alterations of the index and they lead to (22). The correction (22) was verified in particular cases by the "deferred approach to the limit" of Richardson and Gaunt [14].

COMPARISON WITH THE DISTRIBUTION OF SIZES OF TOWNS

Towns and wars are both examples of human aggregation. A referee has asked me to compare them. For simplicity I take the distribution of towns from Lotka's book [10A] in Auerbach's idealized form whereby the town of rank r in a given country has a population n such that

$$nr = A, \tag{1}$$

in which A is a constant, namely the population of the largest town. The desired comparison might conceivably be made by arranging the fatal quarrels in order of rank; but there would be many artificially coincident ranks on account of the rounding off of imperfectly known casualties; moreover the gap in the data prevents the extension of rank from the wars to the murders. It is preferable therefore to leave the fatal quarrels as they are already shown on Figure 1, and to transform Auerbach's law so as to see how towns would appear on a diagram of that type. For the purpose of comparison the same symbols will be used in corresponding meanings. Thus in this section

$$n \text{ is the population of a town, and } \mu \text{ will be } \log_{10} n. \tag{2}$$

Again, to match the fatal quarrels, $\bar{q}$ will here be defined as

$$\bar{q} = \frac{\text{number of towns in a range of population}}{\text{number of integers in that range of population}}. \tag{3}$$

Also, as before, let

$$\phi = \log_{10} \bar{q}. \tag{4}$$

Consider two towns of ranks r', r'', of which r'' is the greater. Let them mark the ends of a range. The number of towns inside the range is strictly $r''-r'-1$. But let us follow the usual statistical practice of regarding half of an end-object as lying on either side of the end. The number of towns in the range is then simply $r''-r'$. Let n' and n'' be the respective populations. With the same convention about ends on the population range, we have from (3)

$$\bar{q} = \frac{\cdot \qquad \cdot}{n' - n''} \cdot \tag{5}$$

Elimination of ranks between (5) and (1) gives

$$\bar{q} = \frac{A}{n'n''} \cdot \tag{6}$$

So that from (4)

$$\phi = \log_{10} A - \log_{10} n' - \log_{10} n''. \tag{7}$$

As for the wars, let μ_0 be the midpoint of the range of μ. It is

$$\mu_0 = \tfrac{1}{2}(\log_{10} n' + \log_{10} n''). \tag{8}$$

From (7) and (8)

$$\phi = \log_{10} A - 2\mu_0, \tag{9}$$

however wide or narrow the range may be. That the constant $\log_{10} A$ should be the same whatever the grouping is a peculiarity of the slope

$$d\phi/d\mu_0 = -2 \tag{10}$$

as shown on page 541, where the correction vanished for $c = 2$.

This slope for the towns is quite close to that of the straight line which in Figure 1 might join the murders to the world wars. Such a broad resemblance between two forms of aggregation is certainly interesting, and may suggest theories. Perhaps the most suggestive formula is

$$q = An^{-2} \tag{11}$$

which follows from (6) when n' is almost equal to n''. Rashevsky [11A], who gives formula (11), has tried-out several explanations.

These overall resemblances do not however conduce to an accurate description of fatal quarrels; for they distract attention from the curvature of the $\phi(\mu)$ graph in Figure 1. The slope of the part relating to wars is certainly not -2. If it were so, then the number of fatal quarrels would increase 10 times for each unit decrease of magnitude. The observed ratio is less than four.

REFERENCES

[1] Anderson, J. G. 1939. *China fights for the world.* Kegan Paul, London, p. 106.

[1A] Carr-Saunders, A. M. 1936. *World Population.* Clarendon Press, Oxford.

[2] Coupland, R. 1939. *The Exploitation of East Africa 1856–1890.* Faber & Faber, London. Chh. VII, XI.

[3] Durbin, E. F. M., and Bowlby, J. 1939. *Personal Aggressiveness and War.* Kegan Paul, London.

[4] Fisher, R. A. 1922. *Phil. Trans. Roy. Soc. Lond.* A 222, 332.

[5] Glover, E. 1933 & 1947. *War, Sadism & Pacifism.* Allen & Unwin, London.

[6] Harding, D. W. 1941. *The Impulse to Dominate.* Allen & Unwin, London.

[7] Hume, M. A. S. 1900. *Modern Spain.* Fisher Unwin, London.

[8] *Japan and Manchoukuo Year Book,* 1938. Tokio.

[9] Jastrzebski, S. de. 1929 article *Death-Rate* in Ency. Brit. XIV edn.

[10] Jeffreys, H. 1939. *Theory of Probability,* p. 148. Clarendon Press, Oxford.

[10A] Lotka, A. J. 1925. *Elements of Physical Biology.* Williams & Wilkins, Baltimore.

[11] Miliukov, P. article *Russia* in *Ency. Brit.* XIV ed., 19, 720 d.

[11A] Rashevsky, N. 1947. *Mathematical Theory of Human Relations.* The Principia Press, Bloomington, Indiana.

[12] Richardson, L. F. in *Nature* of 15 Nov. 1941.

[13] Richardson, L. F. 1945. *J. Roy. Statistical Soc.* 107, 242–250.

[14] Richardson, L. F., and Gaunt, J. A. 1927. *Phil. Trans. Roy. Soc. Lond.* A 226, 299–361.

[15] Sleeman, W. H. 1836. *Ramaseeana.* Military Orphan Press Calcutta, p. 39.

[16] Thrasher, F. M. 1927. *The Gang.* Chicago University Press.

[17] Whitaker's Almanack, annual, London.

[18] Wright, Q. 1942. *A Study of War.* Chicago University Press.

BANK RESERVES AND BUSINESS FLUCTUATIONS*

CLARK WARBURTON
Federal Deposit Insurance Corporation

The belief that severe business fluctuations are caused primarily by the operations of banks, through their influence on the quantity of the circulating medium, was widely held during the nineteenth century and the first three decades of the twentieth. During the latter part of this period variations in bank reserves were believed to be the chief causal factor underneath changes in the circulating medium furnished by the banks. This paper describes a study of factual data since 1918 in the light of that theory. To make the investigation, it was necessary to develop several new time series from available statistical material, and to make appropriate allowance for growth trends. The results support the theory. There is a case by case correspondence between the peaks and troughs in the series of effective member bank reserves and the business cycle reference dates of the National Bureau of Economic Research, with leadership in reserves; and the duration and depth of periods of recession and recovery intervening between the turning points are closely associated with the duration and depth of deviations in reserves from the line of growth.

I WOULD like to begin with two quotations. The first is from Harry E. Miller's book, *Banking Theories in the United States before 1860.*

> "The most common explanation of commercial crises . . . was that they result from the alternating expansion and contraction of the media of payments which occurs wherever a currency of convertible bank notes exists."[1]

The second is from Paul Barnett's book, *Business-Cycle Theory in the United States from 1860 to 1900.*

> "Monetary theories of cyclical changes in business activity received more widespread consideration by writers in the United States from 1860 to 1900 than any other single group of theories."[2]

---

* Talk before the Washington Statistical Society, April 26, 1948.

[1] Harry E. Miller, *Banking Theories in the United States before 1860* (Harvard University Press, 1927), p. 81.

[2] Paul Barnett, *Business-Cycle Theory in the United States, 1860–1900* (University of Chicago Press, 1941), p. 86.

The belief that severe business fluctuations are caused primarily by the operations of banks, through their influence on the quantity of the circulating medium, continued to be widely held during the first three decades of the present century. During this period and the latter part of the nineteenth century an important supplement to the theory was developed, namely, that variations in bank reserves are the chief causal factor underneath changes in the circulating medium furnished by the banks. In addition, the need for continuous increase in the quantity of circulating medium because of increasing population and productivity was given emphasis; the possibility that habits of use of circulating medium may change was recognized; and various aspects of the routes or modes of impact of bank operations were explored.

To describe with precision all phases of this banking theory of business fluctuations cannot be attempted here. My endeavor to provide a concise but reasonably complete statement of the theory has yielded the following six points.

1. Under a fractional reserve system, changes in the amount of commercial bank reserves alter the quantity of deposits and therefore the quantity of assets which such banks may hold; and this is followed by changes in the actual volume of assets and deposits, because banks of deposit have an inherent tendency to expand their operations to the limit placed upon them by law and interbank relationships. Since deposits form the bulk of the circulating medium, changes in deposits dominate changes in the total circulating medium.

2. Changes in the amount of circulating medium, expressed as deviations from a normal rate of increase accompanying growth in population and productivity and adjusted for any measurable trend in monetary habits, alter the equilibrium center of the price system. That is, deficiency in the circulating medium relative to a reasonable rate of growth produces a falling price level because the output of the economy cannot continue to be sold at prevailing prices; while expansion at an excessive rate gives sellers of goods an opportunity to obtain higher prices and the desire for profits produces this result.

3. A falling price level, in view of fixed and sticky costs, narrows profit margins and induces business to reduce output; while a rising price level tends to increase profit margins and therefore stimulates output. In addition, reluctance or inability of some business enterprises to make rapid changes in selling prices leads to unanticipated inventory accumulation when the equilibrium center of prices falls and to diminished inventories when the equilibrium center rises. Such changes in

inventories, as well as changes in prices, induce businessmen to curtail or expand output.

4. The foregoing effects are likely to be accentuated by changes in velocity of the circulating medium. Contraction of the circulating medium is *per se* a shrinking of the cash holdings of business and individuals, and this leads to more hesitant and cautious use of them. On the other hand, rapid expansion of cash balances is followed by a tendency to shorten the period that cash receipts are held before their disbursement.

5. An additional route of impact of changes in bank reserves is through accompanying changes in rates of interest and in the availability of credit for those who wish to borrow. Since those changes may be synchronous with the beginning of a deviation of bank reserves from trend, the effects of such deviations via this route of impact may precede those via the other routes which I have described.

6. Causation sequences are normally accompanied by time lags. Two such lags between changes in bank reserves and sequential business fluctuations are important: (a) that which may occur between the appearance of the forces operating to change the quantity of reserves and the response of commercial banks to those changes; and (b) that which may occur between changes in the quantity of circulating medium and changes in the level of prices and quantity of output. These time lags vary in length. Under some circumstances the total lag may be of negligible duration; under other circumstances as much as or more than a year.

This hypothesis, as I have indicated, was the dominant aspect of business-fluctuation theory in the nineteenth and early part of the twentieth centuries. However, it was far from universally held, and the validity of various parts of it was challenged by some economists. Efforts at testing it by factual analysis were inconclusive because of inadequacy of data. In recent years the hypothesis has been displaced by other theories to such a great extent that few economists have deemed it worthwhile to examine in its light the more abundant data now available. My purpose in offering this paper is to describe briefly the statistical methodology which I have used in such an examination and to state the results.

Scope of statistical analysis

Variations in the value, quantity, and prices of goods and services sold to final purchasers, in the form of deviations from appropriate trends, were deemed to constitute the basic phenomena of business fluctuations. These, together with the quantity of circulating medium (measured by total deposits and currency excluding interbank obligations), habits of use of circulating medium as expressed in velocity, and bank reserves were selected as the principal statistical series to be examined.

The period from 1918 to date was chosen for the study for two reasons. First, since June 1917, required reserves of banks which are members of the Federal Reserve System have consisted exclusively of deposit balances in the Federal Reserve Banks. Second, most of the necessary statistical series were unavailable for prior dates, nor could reasonably reliable series be constructed from the available information.

Quarterly averages were used in the study. On the one hand annual data are not sufficiently sensitive to the variations commonly designated "cyclical"; and on the other, monthly data are more difficult to work with over a period of three decades and are also subject to short-time variations which may be neglected.

Statistical series

Existing statistics did not provide suitable basic series and these had to be constructed. With the exception of bank reserves, the methods of constructing them are described in articles which have been published in economic journals.[3] These series have since been revised, with slight changes in methodology. With respect to bank reserves, reported data for Federal Reserve member banks were adjusted to quarterly averages, and then adjusted successively for changes in percentage reserve requirements, for shifting of deposits among member banks subject to different percentage requirements, and for changes in the proportion of deposits held by member and nonmember banks. These adjustments, together with the resulting series of effective bank reserves, are shown in Table 3, at the end of this article.

Determination of trends

For initial determination of trends in sales of final products and in the quantity and rate of use of circulating medium, the period of greatest stability of output and prices during the three decades was selected.

---

[3] "Normal Production, Income, and Employment, 1945 to 1965," *The Southern Economic Journal*, Vol. XI, No. 3 (January 1945), pp. 219–45; "The Volume of Money and the Price Level Between the World Wars," *The Journal of Political Economy*, Vol. LIII, No. 2 (June 1945), pp. 150–63; and "Quantity and Frequency of Use of Money in the United States, 1919–45," *The Journal of Political Economy*, Vol. LIV, No. 5 (October 1946), pp. 436–50.

At first the seven-year period, 1923–1929, was used, but in view of deviations which occurred in 1929 the period was shortened to the six years, 1923–1928. Percentage growth trends were used, based initially on averages for the terminal years of the period, 1923 and 1928, both years of high production or "full employment" as that phrase is now used.

The trends thus derived were drawn on semi-logarithmic charts of the data for the entire period covered by the study, and were found to be excellent fits for this period. In addition, the trend in quantity of final products sold was compared with an annual index of production covering the period since 1879; and the trend in circuit velocity of deposits and currency was compared with the ratio of national income to deposits and currency by decades for the period since 1799. The apparent trends for the recent and the earlier decades were similar, and their validity as genuine secular trends was deemed to be established.

Several minor variations in the slopes of the trends were examined to see which provided the best theoretical and practical fit. The final selection took into account the fact that the level of prices of final products was lower in 1928 than in 1923 by about 2 per cent, and the principle that the trend values should be consistent with each other when projected forward for a considerable period of time. The slopes chosen were as follows:

Prices of final products: none.
Quantity and value of final products: 3.6 per cent growth per year, or .888 per cent per quarter.
Circuit velocity of deposits and currency: inverse of a growth in deposits and currency relative to expenditures for final products of 1.35 per cent per year, or .336 per cent per quarter.
Quantity of deposits and currency: 5 per cent growth per year, or 1.227 per cent per quarter (when expressed as amounts, the product of the above two trends).
Bank reserves: same as quantity of deposits and currency.

In the final revision, the selected quarterly trend rates were applied to the 25-quarter period from the last quarter of 1922 to the last quarter of 1928, inclusive. Trend values were computed from the quarter in the center of the period.

Deviations from trend

The quarterly statistical series were next compared with trend values.

Points of upward and downward departure from trend, and of maximum deviation, were noted; and also differences in timing with respect to turning points and departures from trend in the various series. In the case of the transition from the prosperity of the 1920's to the great depression, the timing of departures from trend, on the down side, were as follows:

Bank reserves, and deposits and currency—third quarter of 1928
Value of sales of final products—second quarter of 1929
Quantity of sales of final products—fourth quarter of 1929
Prices of final products—first quarter of 1930
Circuit velocity of deposits and currency—third quarter of 1930

This sequence was found to be typical, though not perfectly uniform, and with considerable variation in the length of the lags.

One of the important results is the fact that deviations from trend in circuit velocity—particularly in the case of the beginnings of downswing—lag behind deviations from trend in quantity of deposits and currency, and except for this lag are in the same direction. This result is in conformity with traditional theory that variation in velocity results from and accentuates the effect of variation in the quantity of circulating medium. The assumption which is made in Keynesian and other contemporary theories that idle money is an independent or leading factor in the course of business fluctuations is inconsistent with the observed facts.

Comparison of turning points with business cycle reference dates

The last element in the statistical methodology to be mentioned here was a comparison, for the period 1918–1939, of turning points in bank reserves and in deposits and currency with each other and with reference dates for business cycles. For turning points in bank reserves and in deposits and currency, downswings and upswings relative to trend of one quarter's duration only were neglected. Other turning points were divided into two groups. Those preceded and succeeded by an upswing or downswing of a year or longer were defined as primary turning points; those preceded or succeeded by a downswing or upswing of two or three quarter's duration were termed secondary turning points. The secondary turning points occur in pairs between a pair of primary turning points.

For the period from 1918 to 1939 five pairs of primary turning points and six pairs of secondary turning points occurred in bank reserves. Of the latter, four pairs showed slight amplitude, that is, the difference between the peak and trough expressed as deviations from trend is about one per cent or less; the other two were of considerably greater severity.

TABLE I

TURNING POINTS IN MEMBER BANK RESERVES AND IN DEPOSITS AND CURRENCY COMPARED WITH BUSINESS CYCLE REFERENCE DATES, QUARTERLY DATA, 1918–1939

Turning points	Effective member bank reserves*	Deposits and currency†	Business cycle reference dates‡
Peak	2nd 1918	..	3rd 1918
Trough	4th 1918	..	2nd 1919
Peak	1st 1920	2nd 1920	1st 1920
Trough	4th 1921	1st 1922	3rd 1921
Peak	1st 1923	2nd 1923	2nd 1923
Trough	2nd 1924	2nd 1924	3rd 1924
Peak	4th 1925	4th 1925	3rd 1926
Trough	1st 1927	1st 1927	4th 1927
Peak	2nd 1928	2nd 1928	2nd 1929
Trough	1st 1932	..	Summer 1932
Peak	4th 1932	..	End of 1932
Trough	2nd 1933	2nd 1933	1st 1933
Peak	4th 1935	3rd 1936	2nd 1937
Trough	3rd 1937	2nd 1938	2nd 1938

* Based on deviations from trend, at 5 per cent per year rate of growth, in average member bank reserves adjusted for changes in percentage reserve requirements, shifting of deposits among member banks subject to different percentage requirements, and changes in the proportion of deposits held by member and nonmember banks. Includes all peaks and troughs in this series preceded by upswings and downswings, respectively, of a year or longer, and all peaks and troughs preceded or followed by an upswing or downswing of two or three quarters' duration with an amplitude of 2 per cent or more.

† Based on deviations from trend in total deposits and currency, excluding interbank obligations—with the same criteria as for member bank reserves. In the omitted cases the quarterly estimates of deposits and currency do not show definite peaks and troughs, though changes occurred in the rate of expansion or contraction relative to trend.

‡ Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (1946), 78 and 82. The dates given here include all of the reference cycles recognized by Burns and Mitchell for the period subsequent to January 1, 1918, and also the intermediate cycle of short duration but large amplitude in 1932, which they mention but treat as a double-bottomed trough without assignment of specific reference dates.

The list of turning points in member bank reserves given in Table 1 includes all primary turning points and the two pairs of secondary turning points with considerable amplitude. The business cycle reference dates include all those given for the period since 1918 by Professors Burns and Mitchell of the National Bureau of Economic Research, together with the intermediate cycle of short duration and substantial amplitude in 1932 which they mention but treat as a double-bottomed trough.[4]

It will be noted that there is a case-by-case correspondence between the 14 turning points in member bank reserves during the period from 1917 to 1939 and the same number of National Bureau business cycle reference dates. The peaks and troughs in member bank reserves preceded the business cycle peaks and troughs in ten of the fourteen cases, were simultaneous in two cases, and lagged by one quarter in two cases. However, in the last two cases the specific data are such as to make unwarranted a conclusion that in those cases the recovery in business activity preceded the recovery in reserves.

TABLE 2

RELATIONSHIP OF DEGREE OF CHANGE IN THE EFFECTIVE AMOUNT OF MEMBER BANK RESERVES TO PERIODS OF BUSINESS RECOVERY AND RECESSION, 1918–1941*

Approximate period of expansion or contraction in reserves	Approximate period of recession or recovery
1918–1919, rapid expansion in reserves	1919 and early 1920, business boom
1920–1921, contraction in reserves	Late 1920 and 1921, severe depression
1922, rapid expansion in reserves	1922 and 1923, prosperity
1923, insufficient growth in reserves	1924, minor depression
1924, rapid expansion in reserves	1925, boom
1925, normal growth in reserves	1926, continuation of prosperity
1926, contraction in reserves	1927, depression
1927, rapid expansion in reserves	1928 and part of 1929, prosperity
1928–1930, insufficient growth in reserves	Late 1929 and 1930, depression
1931 and early 1932, drastic contraction in reserves	1932, intensified depression
Summer and autumn of 1932, substantial expansion in reserves	Latter part of 1932 beginning of recovery
Early 1933, drastic contraction in reserves	Same period, renewed depression and collapse of the banking system
Middle 1933 to end of 1935, rapid growth in reserves	Mid-1933 to 1937, continuous business improvement
1936–1937, drastic contraction in reserves	Late 1937 and early 1938, depression
1938–1940, rapid expansion in reserves	Late 1938 to Pearl Harbor, prosperity.

* For description of the concept of effective member bank reserves, see footnote * to table 1.

The relationship of bank reserves to business fluctuations is not limited to the sequence of turning points; the duration and depth of the periods of recovery and recession intervening between the turning points are closely associated with the duration and depth in deviations in effective member bank reserves from the 5 per cent per year line of growth. The sequences are shown in Table 2.

---

[4] Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (National Bureau of Economic Research, 1946), pp. 78 and 82.

TABLE 3

MEMBER BANK RESERVES, ANNUAL AND QUARTERLY AVERAGES, 1918–1947

Year and quarter	Amount in millions of dollars				Relatives for computation of effective amount (1923–1928=1)			Percentage ratio to trend of effective amount of reserves[7]
	Actual amount			Effective amount (1923–28 basis)[3]	Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]	
	Total[1]	Required[2]	Excess					
1918	*1,496*	*1,382*	*114*	*1,437*	*1.00*	*1.158*	*1.112*	*96.3*
i	1,467	1,377	90	1,411	1.00	1.163	1.119	96.3
ii	1,499	1,298	201	1,480	1.00	1.156	1.141	99.9
iii	1,471	1,319	152	1,432	1.00	1.154	1.123	95.4
iv	1,548	1,535	13	1,424	1.00	1.158	1.065	93.7
1919	*1,718*	*1,636*	*82*	*1,600*	*1.00*	*1.138*	*1.059*	*102.1*
i	1,633	1,559	74	1,479	1.00	1.154	1.045	96.2
ii	1,679	1,607	72	1,561	1.00	1.145	1.065	100.3
iii	1,743	1,652	91	1,644	1.00	1.133	1.069	104.3
iv	1,817	1,726	91	1,717	1.00	1.119	1.057	107.6
1920	*1,835*	*1,772*	*63*	*1,743*	*1.00*	*1.092*	*1.037*	*106.0*
i	1,873	1,771	102	1,766	1.00	1.107	1.044	109.3
ii	1,859	1,804	55	1,749	1.00	1.097	1.032	107.0
iii	1,821	1,808	13	1,727	1.00	1.087	1.031	104.4
iv	1,785	1,705	80	1,730	1.00	1.075	1.042	103.3
1921	*1,672*	*1,623*	*49*	*1,656*	*1.00*	*1.060*	*1.049*	*95.8*
i	1,732	1,642	90	1,709	1.00	1.064	1.050	100.8
ii	1,662	1,629	33	1,657	1.00	1.057	1.054	96.5
iii	1,630	1,586	44	1,620	1.00	1.058	1.051	93.2
iv	1,663	1,636	27	1,636	1.00	1.059	1.042	93.0
1922	*1,781*	*1,738*	*43*	*1,718*	*1.00*	*1.052*	*1.016*	*94.7*
i	1,702	1,641	61	1,659	1.00	1.056	1.029	93.1
ii	1,779	1,739	40	1,710	1.00	1.054	1.013	94.9
iii	1,807	1,785	22	1,730	1.00	1.053	1.008	94.8
iv	1,834	1,787	47	1,774	1.00	1.046	1.012	96.0
1923	*1,873*	*1,824*	*49*	*1,873*	*1.00*	*1.019*	*1.019*	*98.3*
i	1,897	1,821	76	1,871	1.00	1.031	1.017	100.1
ii	1,870	1,833	37	1,869	1.00	1.021	1.020	98.7
iii	1,850	1,820	30	1,862	1.00	1.014	1.021	97.2
iv	1,874	1,822	52	1,889	1.00	1.011	1.019	97.4
1924	*2,023*	*1,962*	*61*	*1,989*	*1.00*	*1.019*	*1.003*	*99.5*
i	1,906	1,837	69	1,913	1.00	1.011	1.015	97.5
ii	1,943	1,902	41	1,929	1.00	1.015	1.008	97.1
iii	2,079	1,997	82	2,030	1.00	1.023	.999	100.9
iv	2,162	2,112	50	2,082	1.00	1.027	.989	102.3
1925	*2,166*	*2,130*	*36*	*2,134*	*1.00*	*1.010*	*.995*	*101.7*
i	2,163	2,104	59	2,101	1.00	1.016	.987	101.9
ii	2,132	2,102	30	2,109	1.00	1.007	.996	101.1
iii	2,157	2,115	42	2,140	1.00	1.007	.999	101.3
iv	2,214	2,200	14	2,187	1.00	1.009	.997	102.3
1926	*2,209*	*2,191*	*18*	*2,212*	*1.00*	*.999*	*1.000*	*100.3*
i	2,214	2,197	17	2,201	1.00	1.004	.998	101.7
ii	2,196	2,172	24	2,207	1.00	.996	1.001	100.7
iii	2,208	2,184	24	2,216	1.00	.998	1.002	99.9
iv	2,217	2,210	7	2,222	1.00	.997	.999	99.0

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars				Relatives for computation of effective amount (1923–1928=1)			Percentage ratio to trend of effective amount of reserves[7]
	Actual amount			Effective amount (1923–28 basis)[3]	Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]	
	Total[1]	Required[2]	Excess					
1927	*2,290*	*2,286*	*4*	*2,298*	*1.00*	*.987*	*.991*	*99.3*
i	2,232	2,210	22	2,245	1.00	.990	.996	98.8
ii	2,270	2,266	4	2,284	1.00	.987	.993	99.3
iii	2,291	2,285	6	2,300	1.00	.985	.989	98.8
iv	2,366	2,381	−15	2,362	1.00	.986	.984	100.2
1928	*2,355*	*2,342*	*13*	*2,415*	*1.00*	*.973*	*.997*	*99.4*
i	2,386	2,381	5	2,393	1.00	.984	.987	100.3
ii	2,380	2,368	12	2,444	1.00	.971	.997	101.2
iii	2,304	2,292	12	2,393	1.00	.965	1.002	97.9
iv	2,350	2,362	24	2,428	1.00	.970	1.002	98.1
1929	*2,358*	*2,315*	*43*	*2,472*	*1.00*	*.967*	*1.014*	*96.9*
i	2,360	2,314	47	2,473	1.00	.962	1.008	98.7
ii	2,306	2,270	37	2,452	1.00	.960	1.021	96.7
iii	2,330	2,293	37	2,465	1.00	.964	1.020	96.0
iv	2,434	2,382	52	2,498	1.00	.981	1.007	96.1
1930	*2,378*	*2,323*	*55*	*2,463*	*1.00*	*.965*	*.999*	*91.9*
i	2,328	2,277	51	2,417	1.00	.963	1.000	91.9
ii	2,366	2,319	47	2,456	1.00	.964	1.001	92.2
iii	2,402	2,340	62	2,492	1.00	.963	.999	92.5
iv	2,418	2,357	61	2,488	1.00	.968	.996	91.2
1931	*2,323*	*2,235*	*89*	*2,396*	*1.00*	*.974*	*1.005*	*85.2*
i	2,396	2,321	76	2,448	1.00	.974	.995	88.6
ii	2,389	2,305	84	2,461	1.00	.968	.997	88.0
iii	2,362	2,247	115	2,443	1.00	.973	1.006	86.3
iv	2,148	2,066	82	2,233	1.00	.982	1.021	77.9
1932	*2,114*	*1,858*	*256*	*2,216*	*1.00*	*.984*	*1.033*	*75.0*
i	1,928	1,883	46	2,040	1.00	.976	1.033	70.3
ii	2,065	1,844	221	2,193	1.00	.978	1.039	74.7
iii	2,086	1,812	273	2,203	1.00	.981	1.036	74.1
iv	2,373	1,892	481	2,428	1.00	1.000	1.023	80.7
1933	*2,306*	*1,784*	*522*	*2,287*	*1.00*	*1.029*	*1.021*	*73.7*
i	2,218	1,817	401	2,240	1.00	1.009	1.019	73.6
ii	2,090	1,736	354	2,054	1.00	1.041	1.023	66.6
iii	2,334	1,776	559	2,314	1.00	1.032	1.023	74.2
iv	2,581	1,808	773	2,541	1.00	1.033	1.017	80.4
1934	*3,664*	*2,104*	*1,560*	*3,420*	*1.00*	*1.059*	*.991*	*105.0*
i	2,961	1,918	1,044	2,867	1.00	1.041	1.008	89.7
ii	3,689	2,072	1,616	3,476	1.00	1.056	.995	107.4
iii	3,973	2,164	1,809	3,671	1.00	1.065	.984	112.1
iv	4,034	2,263	1,771	3,666	1.00	1.073	.975	110.6
1935	*4,999*	*2,532*	*2,467*	*4,255*	*1.00*	*1.103*	*.941*	*124.5*
i	4,469	2,358	2,112	3,965	1.00	1.082	.960	118.1
ii	4,731	2,478	2,254	4,072	1.00	1.091	.939	119.8
iii	5,148	2,599	2,550	4,305	1.00	1.112	.930	125.1
iv	5,647	2,693	2,955	4,679	1.00	1.126	.933	134.4

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars					Relatives for computation of effective amount (1923–1928=1)				Percentage ratio to trend of effective amount of reserves[7]	
	Actual amount			Effective amount (1923–28 basis)[3]		Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]			
	Total[1]	Required[2]	Excess								
1936	*5,988*	*3,476*	*2,512*	*4,008*		*1.25*	*1.135*	*.923*		*111.6*	
i	5,669	2,761	2,908	4,666		1.00	1.131	.931		132.4	
ii	5,474	2,840	2,634	4,477		1.00	1.131	.925		125.5	
iii	6,129	3,723	2,406	3,307		1.50	1.137	.920		91.6	
iv	6,681	4,579	2,103	3,583		1.50	1.140	.917		98.0	
1937	*6,829*	*5,603*	*1,226*	*2,910*		*1.94*	*1.119*	*.921*		*77.2*	
i	6,722	4,850	1,872	3,103		1.75	1.134	.916		83.8	
ii	6,878	5,759	1,118	2,817		2.00	1.122	.919		75.2	
iii	6,800	5,958	842	2,820		2.00	1.113	.923		74.3	
iv	6,917	5,845	1,073	2,899		2.00	1.106	.927		75.5	
1938	*7,931*	*5,415*	*2,516*	*3,631*		*1.81*	*1.120*	*.925*		*91.7*	
i	7,246	5,818	1,428	3,043		2.00	1.106	.929		78.3	
ii	7,645	5,192	2,453	3,650		1.75	1.112	.929		92.8	
iii	8,161	5,194	2,967	3,837		1.75	1.123	.924		96.4	
iv	8,673	5,458	3,215	3,993		1.75	1.137	.916		99.1	
1939	*10,344*	*5,957*	*4,387*	*4,571*		*1.75*	*1.159*	*.898*		*100.0*	
i	8,992	5,562	3,430	4,101		1.75	1.139	.909		100.5	
ii	9,902	5,774	4,128	4,449		1.75	1.151	.905		107.7	
iii	10,808	6,072	4,736	4,735		1.75	1.166	.894		113.3	
iv	11,674	6,421	5,253	4,999		1.75	1.181	.885		118.1	
1940	*13,246*	*6,922*	*6,324*	*5,483*		*1.75*	*1.197*	*.868*		*125.6*	
i	12,187	6,579	5,608	5,141		1.75	1.187	.876		120.0	
ii	13,128	6,799	6,329	5,468		1.75	1.195	.871		126.1	
iii	13,595	7,015	6,580	5,595		1.75	1.201	.865		127.4	
iv	14,074	7,294	6,780	5,729		1.75	1.206	.859		128.9	
1941	*13,407*	*8,078*	*5,329*	*5,206*		*1.81*	*1.212*	*.849*		*113.6*	
i	14,105	7,586	6,519	5,679		1.75	1.212	.854		126.2	
ii	13,483	7,796	5,688	5,394		1.75	1.214	.850		118.4	
iii	13,103	7,994	5,109	5,224		1.75	1.214	.847		113.3	
iv	12,936	8,935	4,001	4,525		2.00	1.208	.845		97.0	
1942	*12,649*	*9,976*	*2,673*	*4,413*		*2.00*	*1.187*	*.826*		*91.7*	
i	12,890	9,590	3,301	4,456		2.00	1.215	.840		94.3	
ii	12,594	9,820	2,774	4,293		2.00	1.219	.831		89.8	
iii	12,444	10,182	2,262	4,185		2.00	1.222	.822		86.5	
iv	12,668	10,313	2,355	4,719		2.00	1.090	.812		96.3	
				A	B			A	B	A	B
1943	*12,628*	*11,117*	*1,511*	*5,107*	*4,793*	*2.00*	*1.080*	*.874*	*.820*	*101.1*	*94.9*
i	13,177	11,268	1,909	4,982	4,982	2.00	1.095	.828	.828	100.4	100.4
ii	12,345	10,631	1,713	4,965	4,656	2.00	1.079	.868	.814	98.9	92.7
iii	12,624	11,330	1,293	5,240	4,783	2.00	1.077	.894	.816	103.1	94.1
iv	12,366	11,238	1,128	5,239	4,752	2.00	1.067	.904	.820	101.8	92.4
1944	*13,220*	*12,170*	*1,050*	*6,107*	*5,199*	*2.00*	*1.055*	*.973*	*.829*	*115.1*	*98.0*
i	12,572	11,484	1,087	5,516	4,918	2.00	1.061	.931	.830	105.9	94.4
ii	13,018	12,069	947	5,981	5,036	2.00	1.060	.974	.820	113.4	95.5
iii	13,127	12,051	1,074	6,226	5,206	2.00	1.049	.995	.832	116.7	97.6
iv	14,165	13,075	1,088	6,704	5,636	2.00	1.048	.992	.834	124.1	104.3

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars					Relatives for computation of effective amount (1923–1928=1)				Percentage ratio to trend of effective amount of reserves[7]	
	Actual amount			Effective amount (1923–28 basis)[3]		Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]			
	Total[1]	Required[2]	Excess	A	B			A	B	A	B
1945	*15,050*	*13,928*	*1,122*	*7,340*	*6,121*	*2.00*	*1.032*	*1.006*	*.836*	*131.8*	*109.9*
i	14,173	13,144	1,028	6,807	5,694	2.00	1.038	.997	.834	124.5	104.1
ii	15,064	13,971	1,092	7,321	6,022	2.00	1.038	1.009	.830	132.3	108.8
iii	15,049	13,925	1,122	7,422	6,244	2.00	1.026	1.012	.837	132.5	111.1
iv	15,915	14,673	1,241	7,810	6,543	2.00	1.024	1.005	.842	137.7	115.4
1946	*15,967*	*14,990*	*977*	*7,364*	*6,669*	*2.00*	*1.009*	*.932*	*.844*	*125.9*	*114.2*
i	15,715	14,556	1,157	7,618	6,495	2.00	1.015	.984	.839	132.7	113.1
ii	15,751	14,776	975	7,392	6,580	2.00	1.009	.947	.843	127.2	113.2
iii	16,070	15,168	902	7,285	6,734	2.00	1.007	.913	.844	123.8	114.7
iv	16,331	15,458	873	7,159	6,867	2.00	1.006	.882	.846	120.2	115.5
1947	*16,457*	*15,589*	*868*	*7,043*	*6,989*	*2.00*	*1.001*	*.857*	*.850*	*114.7*	*113.8*
i	16,135	15,293	842	6,962	6,833	2.00	1.000	.863	.847	115.5	113.4
ii	16,021	15,220	801	6,894	6,805	2.00	.997	.858	.847	113.0	111.5
iii	16,565	15,664	901	7,066		2.00	1.001	.854		114.4	
iv	17,107	16,177	930	7,250		2.00	1.004	.851		115.9	

---

[1] From monthly averages of daily figures, *Banking and Monetary Statistics*, pp. 369–70 and 396, and issues of *Federal Reserve Bulletin* 1942–1948.

[2] For 1929–1947, *Banking and Monetary Statistics*, p. 396, and issues of *Federal Reserve Bulletin*, 1942–1948, with amount for first quarter of 1933 estimated in part from weekly data. For 1919–1928 estimated from (a) required reserves on call dates nearest June 30 and December 31, and (b) ratios of average total deposits other than interbank in reporting member banks for quarters immediately preceding and subsequent to June 30 or December 31 to the amount for the call date, with a slight adjustment on the basis of reported average required reserves for 1929. Data from *Banking and Monetary Statistics*, pp. 132–42, 369–71, 373–75, 387, 395, and 396 and issues of *Federal Reserve Bulletin*, 1919.

[3] Average total reserves adjusted by relatives of percentage requirements, deposit shifting, and deposits not requiring reserves (divided by the first two and multiplied by the third). Annual figures are averages of quarterly figures and may differ from those computed from the average annual relatives. For the difference between columns A and B, 1943–1947, see footnote 6.

[4] Ratio of percentage reserve requirements to such requirements during 1923–1928. Change in percentage requirements during a quarter are treated as in force throughout the quarter, in view of the fact that in most instances the changes were announced in advance. Reduction in reserve requirements for central reserve city banks only in 1942 is not taken into consideration here (see footnote 5).

[5] Proportion of net demand plus time deposits in member banks in each of four categories (net demand deposits in central reserve city banks, net demand deposits in reserve city banks, net demand deposits in country banks, and time deposits), weighted by the relative percentage reserve requirements (13, 10, 7 and 3, respectively, until June 1942, and 10, 10, 7 and 3, respectively, subsequent to that date), converted to relatives with the average for 1923–1928=1. These percentages take into consideration the change in reserve requirements in 1942 applicable only to banks in central reserve cities. Deposit figures for 1929–1947 are from monthly or semi-monthly daily averages, and those for 1918–1928 are interpolations from deposits on call report dates (from *Banking and Monetary Statistics*, pp. 73, 81, 87, 93, and 99; annual reports of the Board of Governors of the Federal Reserve System, 1931–1937; and issues of the *Federal Reserve Bulletin*, 1938–1948).

[6] Interpolated from ratio of total deposits adjusted to total net demand plus time deposits in member banks for June 30 and December 31 of each year, converted to relatives with the average for 1923–1928=1.0. Data for total deposits adjusted from *Banking and Monetary Statistics*, pp. 34–35, with figures for December 31, 1917–1922, interpolated, and issues of *Federal Reserve Bulletin*, 1942–1948; data for total net demand plus time deposits in member banks from *Banking and Monetary Statistics*, pp. 73–75, and issues of *Federal Reserve Bulletin*, 1942–1948. For 1943–1947, figures in column A refer to data including, those in column B to data excluding, United States government deposits.

[7] Trend values are computed at 1.227 per cent increase per quarter; and are based on the average effective reserves during the 25-quarter period from the 4th of 1922 to the 4th of 1928, inclusive, centered at the middle quarter of that period.

THE ORDERING OF *n* ITEMS ASSIGNED TO *k* RANK CATEGORIES BY VOTES OF *m* INDIVIDUALS

GARRET L. SCHUYLER
Rear Admiral, USN (Retired)

A method is proposed for arranging in order of acceptability a set of n items (research projects, candidates for positions, articles used, viewed, or smelled) on the basis of judgments of m judges. Unlike previous methods, this permits each judge independently to group in any manner those items tied for his favor, and to omit voting on any items concerning which he cannot decide. The sampling probabilities for estimating accuracy are not here considered.

INTRODUCTION

WE MAY STATE in the following most general terms, the problem faced by a panel of judges ranking items, projects or persons according to some immeasurable quality.

PROBLEM

Items $P_1, P_2, \cdots, P_n$ are to be arranged in order by votes of individuals $M_1, M_2, \cdots, M_m$. Let u_i be number of items concerning which individual M_i is too undecided to express his opinion, and $q_i, r_i, s_i, \cdots$ numbers of items he assigns to categories $a, b, c, \cdots$ of decreasing worthiness. Then upon what criterion C_j is the position of item P_j among the other items best based?

SOLUTION[1]

Consider a table in which each column corresponds to an item and each row to a voter, with letters such as $a, b, c, \cdots$ to indicate the ratings of the several projects, or u for "undecided." The letters $a, b, c, \cdots$ are to be replaced by numbers in the same relative order, and the totals of the columns are to be taken to indicate the order of excellence to be assigned as a composite judgment to the items.

---

[1] M. G. Kendall: The Advanced Theory of Statistics. Griffin, 1943, p. 410.

For this purpose the first rank numbers to suggest themselves are the midranks

$$(1/2)(q_i + 1),\ q_i + (1/2)(r_i + 1),\ q_i + r_i + (1/2)(s_i + 1),\ \cdots$$

in the blocks of items receiving a-votes, b-votes, c-votes, . . . respectively of member M_i. But the average of the $(n-u_i)$ rank numbers which M_i assigns on this basis is only $(1/2)(n+1-u_i)$; and this function of the number of u-votes varies from member to member.

To have rank numbers from all members' votes properly additive, the average rank numbers for all members' votes should be equal. To effect this it seems simplest to multiply each member's midrank numbers by an adjustment factor

$$A_i = \frac{n+1}{n+1-u_i} \tag{1}$$

which will make the average rank numbers for a-votes, b-votes, etc. of each member equal to $(1/2)(n+1)$. Also, a rank number $(1/2)(n+1)$ should be assigned for each u-vote.

If R_{ai}, R_{bi}, R_{ci}, etc. and R_{ui} are the rank numbers thus assigned to items receiving respectively a-votes, b-votes, c-votes, etc. and u-votes of member M_i, we thus have

$$R_{ai} = A_i(1/2)(q_i + 1),$$

$$R_{bi} = A_i(q_i + (1/2)(r_i + 1)),$$

$$R_{ci} = A_i(q_i + r_i + (1/2)(s_i + 1)),$$

$$R_{ui} = (1/2)(n + 1).$$

The criterion for item P_j is then the sum C_j of the rank numbers R of the votes of all members on this item.

FLEXIBILITY OF THE METHOD

These equations do not restrict panel members to the use of any particular number of categories or ranks, nor do they even require that all members use the same series of ranks in the same voting. A member may or may not leave the ratings of individual items undecided.

The minimum possible number k of different ranks which a member may meaningfully assign is obviously 2. When all members use the maximum value n of k, the order of the C_j is the same as that indicated by Kendall[2] without consideration of the more general problem here described.

THE CASE $k=3$

Principal interest seems most likely to lie, however, in less idealized cases where, unlike judges encountered in the textbook examples, panel members will not

(a) agree to vote on every item (i.e., to use no u-votes), nor

(b) agree to avoid tie votes (i.e., to use as many ranks as there are items).

Recently, for example, panels of 15–20 voters were required to arrange series of 40–122 problems in order of worthiness for retention on a possibly reduced research budget. Few members had specialized on all parts of the field covered. Nearly every member encountered items on which he felt himself insufficiently well qualified to cast a possibly decisive vote.

It was soon agreed that members might withhold opinions on any items with which they felt themselves too unfamiliar; and their comments showed certain members of some panels would cast only u-votes on a majority of the items.

Again very unlike those most obliging judges encountered in textbook examples, these panel members declined to use dozens of ranks. The most they could be persuaded to use was only three ranks (corresponding perhaps to rough classifications of "very worthy," "worthy" "less worthy").

As a practical illustration consider Table I in the first four blocks of which are shown (a) identifying numbers of voters, (b) their votes on two very closely competing items P_{33} and P_{28}, (c) their numbers of votes on all items, and (d) rank numbers, and their averages, for items P_{33} and P_{28}.

CRITERIA WHEN $k=3$

In spite of its comparative neglect in textbooks this case in which no voter discriminates more than three ranks is sufficiently common to have developed an approximate or "1, 2, 3 method" for its treatment.

---

[2] For completeness one must mention also the possible, though I think less promising methods, by which the order of items (a) is so selected as to minimize the number of "inversions" of rank order in relation to ranks assigned by the various members' votes or (b) is taken in accordance with *net* numbers of votes which rank each item superior to other items or (c) is based upon "Scores for Ranked Data," Table XX, Fisher and Yates' "Statistical Tables for Biological, Agricultural, and Medical Research" (1943).

TABLE I

ACTUAL DATA OF VOTES ON ITEMS P_{32} AND P_{28} OF A 41-ITEM SERIES VOTED ON BY A 15-MEMBER PANEL

M	Votes on Items		All Votes				Rank No's. By C_j/n		"Scores"		Rank No's. By C_j'/n	
	P_{32}	P_{28}	a	b	c	u	P_{32}	P_{28}	P_{32}	P_{28}	P_{32}	P_{28}
1	c	c	16	9	16		33.5	33.5	3	3	34.7	34.7
2	c	b	11	16	13	1	35	20	3	2	34.7	21
3	b	b	16	16	9		24.5	24.5	2	2	21	21
4	b	b	9	22	10		20.5	20.5	2	2	21	21
5	c	u	9	13	9	10	35.4	21	3	—	34.7	—
6	b	b	22	11	8		28	28	2	2	21	21
7	b	c	22	13	6		29	38.5	2	3	21	34.7
8	b	c	11	17	13		20	35	2	3	21	34.7
9	c	b	15	15	11		36	23	3	2	34.7	21
10	c	c	13	13	15		34	34	3	3	34.7	34.7
11	a	b	17	7	17		9	21	1	2	7.3	21
12	c	b	15	16	10		36.5	23.5	3	2	34.7	21
13	c	c	15	12	14		34.5	34.5	3	3	34.7	34.7
14	b	b	18	14	9		25.5	25.5	2	2	21	21
15	b	c	24	9	8		29	37.5	2	3	21	34.7
			Means				Means		Means			
			15.5	13.5	11.2	.7	28.7	28.0	2.40	2.43	26.4	26.7
		Max	24	22	17	10	±5.9	±5.9	±.63	±.49	±7.6	±6.7
		Min	9	7	6	0	(m.d.)		(m.d.)		(m.d.)	

Votes

Item P_{32}	1	7	7	0
Item P_{28}	0	8	6	1

Example: For judge M_5

$$A_5 = \frac{41+1}{41+1-10} = 1.31$$

$$R_{c5} = 1.31(9+13+\tfrac{1}{2}(9+1)) = 35.4$$

$$R_{u5} = \tfrac{1}{2}(41+1) = 21$$

With this the a-votes, b-votes and c-votes are respectively assigned the arbitrary weights 1, 2, 3 and items are ranked according to weighted averages.

When a_j, b_j, c_j are numbers of a-votes, b-votes, and c-votes respectively which item P_j receives from all members, the criterion often used is

$$C_j' = \frac{a_j + 2b_j + 3c_j}{a_j + b_j + c_j}.$$

This gives an approximation to the order specified by the C_j in that, if each voter places exactly one-third of the items in each category, the order among the numbers C_j' is the same as that among the C_j. In other cases the order may differ.

The ordering of the items according to C_j' has the dubious property of making a c-vote and an a-vote count always the same as two b-votes for an item. Furthermore, this method makes a u-vote count the same as the average of all other votes on the *same* item, whereas ordering by the C_j makes a u-vote count the same as the average of the votes on *all* items.

With moderately large numbers of voters these differences cancel out surprisingly well. For example, in the voting on 41 items by a 15-member panel cited above, there was only one pair of items whose relative rank differed by the two methods. These were the items P_{28} and P_{33} for which the detailed vote is shown in Table I. Even in this single case the margin can doubtless be ascribed to chance.

CONCLUSIONS

The simple and rough ordering according to C_j', carried out to the nearest tenth of a unit, and with differences of less than two-tenths of a unit between pairs of items disregarded, will probably meet most accuracy requirements encountered in practice, at least if the proportion of undecided votes is as low as in Table I. But the better method using C_j involves only a negligible amount of additional labor and, if only because of its extreme generality and flexibility, it is believed that it should be more widely understood.

The degree of consistency in data of this sort may be examined in various ways, and probability calculations are essential in examining it. Interest in this aspect of the problem, and the appropriate methods, may vary widely according as the panel is intended (a) to furnish a sample of the opinion in some very large and not too select, population, or (b) to furnish a record only of its own collective judgments.

LEVELS OF SIGNIFICANCE FOR VARIANCE RATIO OF TWO SAMPLES OF EQUAL SIZE

C. J. KIRCHEN*

University of Michigan Aeronautical Research Center, Ypsilanti, Michigan

THIS CHART permits two-tailed probability evaluations of ratios of variances of two equal-sized samples at other than the conventional points of 5 and 1 per cent which are 10 and 2 per cent two-tailed levels, respectively. The chart was plotted from the tables of Thompson and Merrington (*Biometrika:* 33, pp. 73–88), the 20 per cent table of H. W. Norton (*Statistical Tables for Biological, Agricultural and Medical Research*, by Fisher and Yates; Oliver and Boyd, Ltd., Edinburgh and London), the .1 per cent table of Colcord and Deming (*Sankhya*, ii, pp. 423–424). The plotted points for 19 and 49 degrees of freedom were obtained by harmonic interpolation on n_1, linear interpolation on n_2.

The reader who is interested in acquainting himself with the mathematical basis for this chart is referred to *Introduction to Mathematical Statistics* by P. G. Hoel; John Wiley and Sons, Inc., New York; 1947; pp. 152–154.

---

* Formerly with Remington Arms Company, Inc., Bridgeport, Conn.

Corrections to article published in Volume 43, No. 242, June 1948

ON ESTIMATING PRECISION OF MEASURING INSTRUMENTS AND PRODUCTS VARIABILITY

FRANK E. GRUBBS

Page 244, middle of page: Change $x_1, x_2, \cdots x_1 \cdots x_n$ to read $x_1, x_2, \cdots x_i, \cdots x_n$

Page 245, second line from bottom: Change $x_i+\bar{x}$ in second term to read $x_i-\bar{x}$

Page 245, bottom line: change c_1 to c_i.

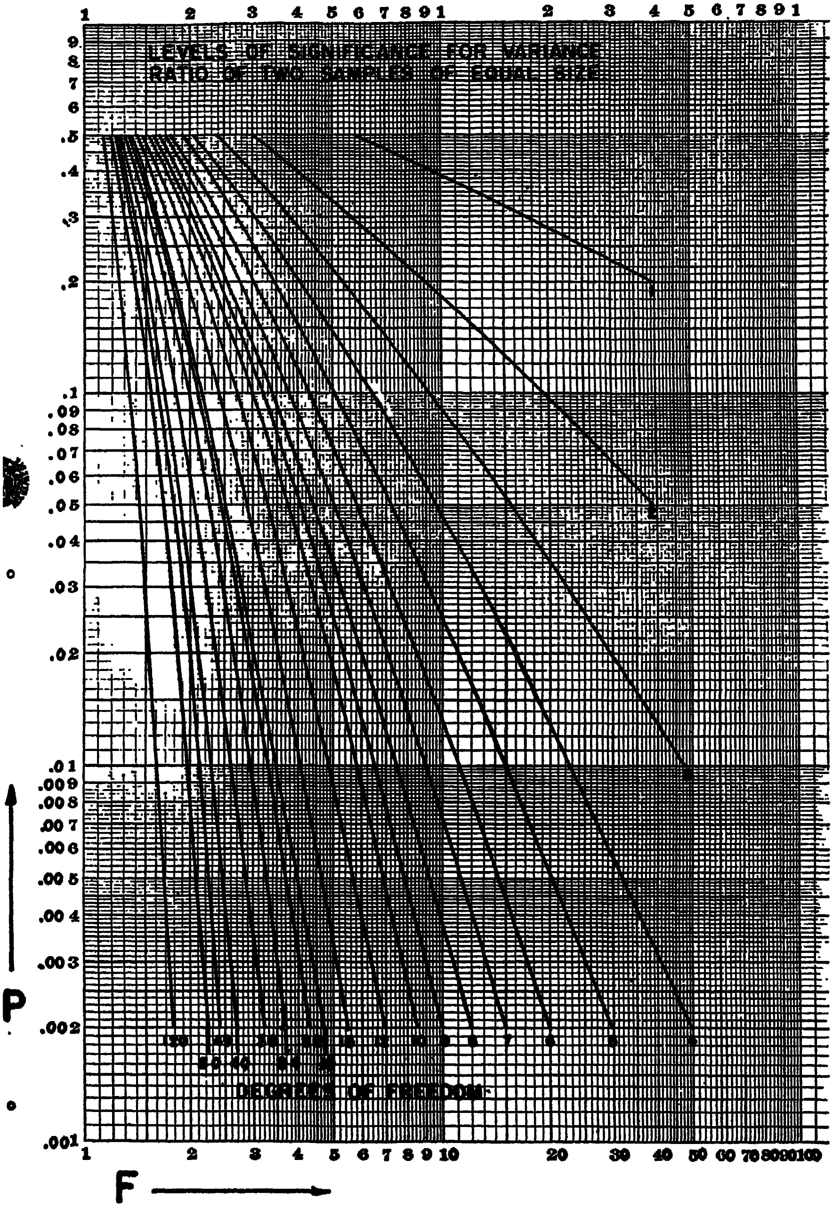
Page 249, line 8: Change x_1 to x_i and change x_{i1} to x_i

Page 253, tenth line from bottom of page: Change $e_{11}+e_{12}$ to $e_{11}-e_{12}$

Page 261, formula (30): Change e_n to e_N

Page 264, line 17: Change the word "variance" to read "variation"

Page 264, footnote, third line from bottom: Change "use" to read



The author wishes to acknowledge with thanks the permission granted by the Remington Arms Co. to publish the attached chart, which was devised in connection with work in which the author was engaged with the company.

MAIN EFFECTS AND INTERACTIONS

D. J. Finney
University of Oxford

When the effects of two factors are simply additive, choice of estimates can be made entirely on grounds of minimal variance. When interaction exists, estimation of the magnitudes of main effects and interactions requires an arbitrary definition of what constitutes a main effect. Whether the class frequencies be equal, proportionate, or disproportionate, the definition must conform to the use that is to be made of the results and must not be dependent merely on the convenience of an analysis of variance.

The analysis of variance is now so widely recognized as one of the most powerful techniques for the examination of statistical data that its users are sometimes in danger of regarding it as an end in itself instead of as a means to an end. In a well designed experiment, for example, the analysis of variance of the results and the summarizing of the effects of treatments are closely related and may even be merely different aspects of the same set of calculations; nevertheless, few things betray the inexperienced statistician more readily than a triumphant presentation of an elaborate analysis of variance table coupled with an almost complete neglect of any statement of treatment means. With fortuitous or observational data, in which symmetry may be less marked than in data from a planned experiment, analysis of variance is still a valuable method of studying different components of variation, but additional care is needed in order to ensure that the pattern of analysis adopted is appropriate to the data and to proceed satisfactorily from analysis to summary of results.

In a recent letter to *Nature*, Vajda [1] has drawn attention to a difficulty that occurs in the definition of a "main effect" for an analysis of variance of data with unequal sub-class frequencies. Proper understanding of his point is important even with data having equal frequencies in the sub-classes, for only when the arbitrary nature of any definition of a main effect, at least if interactions are present, is appreciated can the analysis of variance and the summarizing of conclusions be seen in right perspective. Yates [2] has corrected certain misconceptions that might arise from an incautious reading of Vajda's letter. A fuller statement, however, may be valuable to some whose knowledge of the analysis of variance is based more on text books of statistics than on their own experience.

The problem may be illustrated by reference to data on the interaction of two factors. Suppose that n_{ij} observations are available on a variate x, at a level i of factor A and a level j of factor B in combination, and that the mean yield (using "yield" as a general term for the variate under analysis) for this combination A_iB_j is x_{ij}. Suppose further that p levels of A and q levels of B have been examined; if certain combinations of levels are lacking, some of the n_{ij} will be zero. The mean yields may be set out as a $p\times q$ table:

		Level of A				
	$j \diagdown i$	1	2	3	$\cdots$	p
	1	x_{11}	x_{21}	x_{31}	$\cdots$	x_{p1}
Level of B	2	x_{12}	x_{22}	x_{32}	$\cdots$	x_{p2}
	3	x_{13}	x_{23}	x_{33}	$\cdots$	x_{p3}
		x_{1q}	x_{2q}	x_{3q}	$\cdots$	x_{pq}

and the sum of squares of deviations from the general mean may be partitioned as follows:

	d.f.
Between classes	$pq - 1$
Within classes	$\underset{i}{S}\underset{j}{S}\, n_{ij} - pq$
Total	$\underset{i}{S}\underset{j}{S}\, n_{ij} - 1$

where $\underset{i}{S}$, $\underset{j}{S}$ indicate summations over all values of the suffixes i, j, respectively.[1] This much is easy: difficulties arise only when attempts are made to examine particular components of the sum of squares between classes, and especially those components usually described as "main effects."

The effect of factor A is said to be independent of factor B if the difference in mean yield between any two levels of A is the same at each level of B, except for variations attributable to random sampling: $(x_{ij}-x_{i'j})$ is dependent upon i, i' but, apart from sampling variation, independent of j. That differences in mean yields between any two levels of B are similarly independent of the level of A necessarily follows. The *main effect* of A is then naturally defined as the set of differences between levels of A, and will be estimated by fitting a set of $(p-1)$ constants for A simultaneously with a set of $(q-1)$ constants for B, using the method of least squares (Yates [4]) and weighting each x_{ij} with n_{ij}.

---

[1] The "total" and "between classes" degrees of freedom will be reduced by 1 for every zero value of n_{ij}.

In a planned experiment, n_{ij} will often be a constant, and the subdivision of the sum of squares between classes into main effects and interaction is then very simple. (Further components for "blocks" or other constraints may have to be removed from what has been described above as "within classes," before a residual error mean square can be formed, but that does not affect the principles discussed here.) The required main effects are the differences between simple arithmetic means of columns and rows in the x_{ij} table. When n_{ij} is not constant, in general the analysis is more complicated, but the particular case of *proportional frequencies*, mentioned by Vajda, can be treated simply. If

$$n_{ij} = u_i v_j,$$

where u_i is independent of j and v_j is independent of i, differences between column-means weighted according to frequencies, namely

$$x_{i.} = \underset{j}{S}\, n_{ij} x_{ij} / \underset{j}{S}\, n_{ij} = \underset{j}{S}\, v_j x_{ij} / \underset{j}{S}\, v_j$$

are easily seen to be orthogonal with the similar set of contrasts based on row means

$$x_{.j} = \underset{i}{S}\, n_{ij} x_{ij} / \underset{i}{S}\, n_{ij} = \underset{i}{S}\, u_i x_{ij} / \underset{i}{S}\, u_i.$$

Furthermore, in virtue of the well-known property of a mean of observations weighted inversely as their variance, differences between the $x_{i.}$ estimate the main effect of A with lesser variance than would any other method of averaging the columns. Similarly the $x_{.j}$ give the most efficient estimate of the main effect of B. It is therefore natural to regard the $x_{i.}$ and the $x_{.j}$ as providing the best estimates of the main effects from the data. The least squares procedure of fitting constants leads to exactly the same results. An analysis of variance, using the familiar formulae to give sums of squares for main effects, leaves a third constituent of the total sum of squares between classes, with $(p-1)(q-1)$ degrees of freedom, which is appropriate to a test of significance of the null hypothesis of independent action of A and B;

this is usually described as a significance test on the interaction, $A.B$.

The difficulty to which Vajda draws attention arises only when the null hypothesis of independence has to be abandoned. The effects of the two factors are then no longer simply additive: the definition of the main effect of A, and in particular of the difference between mean yields for levels i and i', must therefore be re-examined. This difference now varies in some manner dependent upon the level of B, and, even though $n_{ij} = u_i v_j$, considerations other than minimal variance may have to determine the estimated average value for the data. To use $x_{i.}$ as before is tempting, on account of its computational convenience and analytical elegance, for it still leads to the weighted average of quantities $(x_{ij} - x_{i'j})$ having minimal variance and it still gives a neat analysis of variance table. This is the estimation termed by Vajda that of "proportionate frequencies"; he appears to prefer to use mean values for levels of A and B defined by

$$\bar{x}_{i.} = \underset{j}{S}\, x_{ij}/q, \qquad \bar{x}_{.j} = S\, x_{ij}/p,$$

the simple averages of class means, identical with the $x_{i.}$ and $x_{.j}$ only if n_{ij} is the same for all i, j. This he speaks of as the "method of fitting constants," but he uses the phrase in a sense rather different from Yates, who reserved it for the estimation of additive main effects when interactions were ignored. Vajda's usage implies the "estimation" by least squares of $(pq-1)$ parameters from the $(pq-1)$ differences between class means x_{ij}, so giving a perfect fit.

Yates [5] points out that "Inasmuch as the frequencies actually obtained are likely to be similar to those generally occurring in the type of material under consideration, the definition of the main effects implied by the use of proportionate frequencies is usually more appropriate in practice than that based on equal frequencies." Whether in any one instance the definition is "more appropriate" depends upon the particular circumstances, and elegance of analysis alone must not be the criterion. Unless the v_j are proportional to frequencies for the q levels of B specially interesting to the experimenter (as when they are relative frequencies in a natural population of which the experimental material is intended to be representative), there is no good reason for preferring $x_{i.}$ to $\bar{x}_{i.}$. The point to be emphasized is that both are arbitrary: $\bar{x}_{i.}$ arbitrarily assigns equal importance to each of the q values of $(x_{ij} - x_{i'j})$, whereas $x_{i.}$ arbitrarily weights these differences in proportion to v_j. Conceivably, though rarely, the mean values of A required in summarizing the data would be defined by some extraneous system of weighting for levels of B, using weights v_j^* independent of i, but different from v_j, to give

$$x_{i.}^* = \underset{j}{S}\, v_j^* x_{ij} / \underset{j}{S}\, v_j^*.$$

These, and a similar set of $x_{.j}^*$ for B, would estimate two sets of main effects mutually orthogonal and themselves defining the interaction as the residual set of $(p-1)(q-1)$ contrasts. As a method of summarizing the results, this would be quite legitimate, though it would not correspond to a simple analysis of variance. Vajda rightly stresses that "the definition of an effect or of an interaction should not depend upon whether the frequencies are noticed to be such that a change in the definition leads to easier computations," but mistakenly concludes that "the method of fitting constants should therefore always be accepted as the only legitimate one." When the interaction is not negligible, estimation of the main effect of A must depend upon circumstances, and no definition, whether $x_{i.}$, $\bar{x}_{i.}$, or the more general $x_{i.}^*$, can be claimed as "the only legitimate one." The problem forces itself upon the attention for the case of proportionate sub-class frequencies because alternative procedures are then very evident. It still exists at the extremes of symmetry and asymmetry. The wise investigator will take account of this, when possible, by choosing his sub-class frequencies in accordance with the proportionate frequencies condition, so as to combine maximal precision of estimation with the weighting system appropriate to his interests.

When all the n_{ij} are equal, as in the majority of planned experiments, the assumption that the simple average for each level of A over all levels of B gives the ideal estimate of the main effect of A is sometimes made too readily; if there is an interaction between A and B, the interpretation of these averages needs to be most carefully considered. A common situation is that in which the "levels" of factor B represent different methods of performing some operation. If the effect of A depends upon which method is chosen, no type of average A effect can be of much interest unless the q methods tested in the experiment either are the only possibilities or are in some way representative of all possible methods. An A effect averaged over an entirely arbitrary selection of levels of B may give a formal completeness to a summary table of results, but is then meaningless or even misleading: the results should instead be interpreted in terms of the effect of A for each B_j separately. For example, in a field experiment on a root crop, factor A might represent amounts of nitrogenous fertilizer and factor B different forms of cultivation (inevitably a small selection only from all the methods of cultivation that might interest the experimenter). If analysis shows the hypothesis of simple additive effects to be untenable, a nitrogen "main effect" averaged over several cultivation treatments may be little use except as a broad qualitative expression of the results. Any detailed interpretation of the experiment will need to examine either responses to nitrogen under each system of cultivation separately or differential cultivation effects at each level of the fertilizer separately.

When $n_{ij} \neq u_i v_j$, no simple analysis of variance exists to tempt the experimenter to sacrifice meaning for elegance, but again he must not assume that the unweighted averages $\bar{x}_{i.}$ and $\bar{x}_{.j}$ are necessarily what he requires. In the absence of interaction, the estimation problem is always simple in principle, and choice between alternatives can be based on their relative precisions. In the presence of interaction, the definitions of main effects and interactions are inextricably bound together; exact meaning can be given to interaction only as the departure from a specified system of main effects, and definition of the estimates of the latter must involve, explicitly or implicitly, arbitrary weighting of the effects of one factor at different levels of another.

The above discussion has been restricted to two factors and their interaction, but the generalization to a greater number should be immediately apparent. This note is concerned only with the logical basis for speaking of main effects and interactions; the technique of estimating them from non-orthogonal data, and of fitting them into the framework of analysis of variance, has been described by Yates [3, 4] and, in great detail and with wealth of illustrative examples, by Snedecor [1].

REFERENCES

[1] Snedecor, G. W., *Statistical Methods* (4th edition), Iowa State College Press, Ames, Iowa, 1946.

[2] Vajda, S., "Technique of the Analysis of Variance," *Nature*, London, Vol. 160, 27, 1947.

[3] Yates, F., "The Principles of Orthogonality and Confounding in Replicated Experiments," *Journal of Agricultural Science*, Vol. 23, 108–145, 1933.

[4] Yates, F., "The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes," *Journal of the American Statistical Association*, Vol. 29, 51–66, 1934.

[5] Yates, F., "Technique of the Analysis of Variance," *Nature*, London, Vol. 160, 472–473, 1947.

A TEST FOR SYMMETRY IN CONTINGENCY TABLES

ALBERT H. BOWKER
Stanford University

SUPPOSE THAT a sample of N individuals is classified according to two arguments, A and B, in an m by m contingency table and suppose that the categories of classification A are associated in some unique way with the categories of classification B. For example, in the study in which this problem arose the categories were related to the position of plates of stickleback (Gasterosteus) and the two arguments corresponded to right and left sides; it was logical to associate identical patterns on opposite sides. Denote by P_{ij} the probability that a randomly chosen individual belongs to the ith row and jth column of the table. As usual, a dot indicates summation over the replaced index.

DISTRIBUTION OF POSSIBLE EVENTS

	Category	Classification A						Marginal Distribution
		1	2	—	—	—	m	
Classification B	1	P_{11}	P_{12}	—	—	—	P_{1m}	$P_{1.}$
	2	P_{21}	P_{22}				P_{2m}	$P_{2.}$
	—	—	—				—	
	—	—	—				—	
	—	—	—				—	
	—	—	—				—	
	m	P_{m1}	P_{m2}				P_{mm}	$P_{m.}$
Marginal Distribution		$P_{.1}$	$P_{.2}$				$P_{.m}$	

It is desired to test the hypothesis that the above table is symmetric (i.e., $P_{ij}=P_{ji}$). If this hypothesis is true, then $P_{k.}=P_{.k}$; that is, the marginal distribution of individuals is the same for classifications A and B. Further, the conditional distribution with respect to A for the kth category of B is equal to the conditional distribution with respect to B for the kth category of A. This equality of conditional distributions in general implies a stronger equivalence between the classifications than equality of marginal distributions, although for $m=2$ equality of marginal distributions implies complete symmetry. The weaker hypothesis of equality of marginal distributions would also be of interest, especially in the absence of symmetry; this problem appears to be somewhat more difficult to handle by straightforward methods.

The proposed test for the hypothesis of symmetry is of the familiar χ^2 type. Let n_{ij} be the observed frequency in the ith row and jth column of a table of observed quantities arranged as above. Then

$$(1) \qquad \sum_{i>j}\sum \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}}$$

has for large N the Chi-square distribution with $m(m-1)/2$ degrees of freedom. A large value of χ^2 leads to rejection of the hypothesis. Note that the value of χ^2 is computed from non-diagonal frequencies, which is reasonable since diagonal frequencies should not afford evidence tending to disprove the hypothesis of symmetry.

In cases in which the observed frequencies are small and no regrouping of frequencies is possible (e.g., in a 2×2 table) an exact test may be of interest. Specifically, in a 2×2 table an exact test may be made by testing with any table of the binomial distribution the hypothesis that n_{12} and n_{21} are observed "successes" and "failures" in a random sample of $n_{12}+n_{21}$ events. From equation (1) we obtain for a 2×2 table the following χ^2 statistic with one degree of freedom:

$$(2) \qquad \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}}.$$

This is equivalent to the value obtained by a test of goodness of fit of a binomial with $P = \frac{1}{2}$. Formula (2) was obtained by McNemar [1], who lists applications in psychological testing. A problem in the same field prompted a recent query in *The American Statistician* [2].

Several intuitive derivations of the above test may be given, but a rigorous justification may be obtained from a theorem of Cramér [3], page 426, which puts the use of χ^2 tests in contingency tables on a sound theoretical basis. This procedure affords an interesting example of a routine application of the general theory and might well be an exercise for a course in the theory of χ^2 tests. After verifying certain regularity conditions, the method of Cramér can be applied by computing the quantity

$$(3) \qquad \sum\sum \frac{(n_{ij} - NP_{ij})^2}{NP_{ij}}$$

where the P_{ij} are replaced by their maximum likelihood estimates under the hypothesis of symmetry. These estimates turn out to be $n_{ij}+n_{ji}/2N$ which yield equation (1) on substitution into (3).

An alternate derivation is available if (1) is regarded as a sum of $m\ (m-1)/2$ terms of type (2). This approach requires the assumption that we are sampling from a population with fixed sums for all symmetric pairs of cells. While both methods lead to the same result, the more general one based on Cramér's theorem seems slightly preferable. Until the test procedure is derived, it is not obvious that the use of the conditional distribution with fixed totals for symmetric pairs entails no loss of efficiency.

The problem which prompted this note was suggested by Professor Frank Weymouth.

REFERENCES

[1] McNemar, Quinn, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, Vol. 12, No. 2, 1947, pp. 153–157.

[2] "Question 1," *American Statistician*, Vol. 1, No. 2, 1947, p. 12.

[3] Cramér, Harald, *Mathematical Methods of Statistics*, Princeton University press, 1946.

THE WAR PRODUCTION BOARD'S STATISTICAL REPORTING EXPERIENCE

PART IV

QUESTIONNAIRE UNITS OF MEASURE AND NOMENCLATURE

DAVID NOVICK
University of Puerto Rico
AND
GEORGE A. STEINER
University of Illinois

Effective wartime administration of the industrial system required the use of a common factual language with which government and industry could communicate. No such system existed in the prewar period. An effort to develop a suitable quantitive language was begun early in the war, and met with some of the most difficult problems encountered by wartime management. Although substantial progress was made, a stabilized and commonly understood system of communication was never attained. This article discusses some of the outstanding problems in connection with questionnaire units of measure and terminology. An analysis of the development of material and product classification and terminology is followed by a discussion of the methods employed to standardize concepts relating to general terminology in questionnaire instructions, to units of measure, to reporting time periods and to end-product usage.

PRODUCT NOMENCLATURE AND CLASSIFICATION

THE PROBLEM of developing uniform identification and classification of materials and products was more than a question of simple definition. Standard definitions of both products and materials were available. So also were classification systems. The problem lay in widely varying practices of industry relating to record-keeping methods and definitions, procurement patterns, and commonly used day-to-day expressions. Into this traditional lack of uniformity were injected rapidly changing wartime influences, new products and materials, and substantial modifications and revisions of old products. These conditions lay at the base of the broad problem of unifying product and material concepts.

Product names and other words used to identify them early in the war meant various things to different minds and by no means did they always carry either the dictionary meaning or the technical one given to them in industry. Under the operation of the early preference rating system, products and materials for which preferred treatment was extended were defined in broad terms, such as bombs, cranes, explosives, and tin. It was found very soon that manufacturers in reporting like products under this system used so many different names and so many varying classifications that identification for tabulation, allocation, and issuance of preference ratings became exceedingly difficult.

As supply relative to demand tightened for an increasingly large number of products and materials, and as administrators became more familiar with their products and the problems of controlling them, a drive began for greater precision of definition. But getting more detail intensified the problem of nomenclature since the result was often a further entanglement in varying industry usage and variations of a given product in grade, size, weight, cost and other characteristics, such as horsepower in the case of certain machines, warmth-conserving qualities in the case of waterfowl feathers, or wear-resistance in the case of leather and textiles.

Relatively little improvement was noted in more precisely defining products and materials given in the *Priorities Critical List* to which preferred treatment could be extended on rating certificates. The move for more precision really began with the introduction of blanket ratings authorized by P Orders in the spring of 1941. The earliest P Orders were couched in vague terms with respect both to the product to which preference ratings might be applied and to the materials and components to which ratings might be extended. Continuous and often voluminous successive amendments, however, brought greater and greater detail into terminology.

Order P-2 illustrates this evolution. Dated March 26, 1941, this order related vaguely to "material for the production of machine tools." No further definition of machine tools was given. But the list of 11 materials to which rating could be extended to facilitate machine tool production with "Machine Tools, consisting of power-driven metal working machines, hydraulic and mechanical presses, and welders," and continued with components such as "motors and other electric equipment," "abrasives," "machine parts and equipment," etc. In the next month a series of amendments more precisely defined equipment, viz., "power-driven metal-working machines and testing equipment," "drill presses," "tornio and automatic screw machines," "nibblers and cutoff machinery," "magnetic chucks and grinding machines," and so on. With each successive amendment the list of specific tools became longer and more precise.

Reporting forms naturally followed the trends established by the orders to which they were attached. Early reporting forms provided blank spaces for manufacturers to report products coincident with the move for more precise identification. Later forms often preprinted on the form itself the exact products to be reported. This development, obviously, served not only to limit reporting to products specified but also facilitated uniformity of response and understanding.

Limitation orders, first issued in mid-1941, generally described the products to which they applied in far more precise terms than early P Orders. This was due partly to experience in connection with P Orders and partly because the L Orders, as an instrument for limiting production of a specified item, had to be precise to insure compliance. In addition, of course, the limiting of production necessitated considerable care because it moved deeply into the entire problem of industrial profits and conversion to war work.

The definition of trucks and motor vehicles set forth in L-1, issued in August 1941, did not change materially in successive amendments to the order or to other L Orders issued to apply to model types originally incorporated in L-1. L-26, which was the next to last L Order issued in 1941, restricted the production of farm machinery, farm equipment, and parts. Products covered by the order were listed in great detail. The reporting instrument first used under the order, PD-284, preprinted each of the approximately 175 specific pieces of machinery and equipment on the stub of the form. This sort of pattern was followed by many L Orders and related forms, such as L-7 for refrigerators and its forms PD-176, 177, and 417; L-17 for vacuum cleaners and its forms PD-170, 171 and 417; and L-21 for automatic phonographs and its related forms PD-182 and PD-417.

Experience of administration brought greater improvement in product definition. But among the many orders, regulations and questionnaires, there existed great variation in definitions, a situation which although constantly improved lasted throughout the war period. Attempts of the WPB Requirements Committee and the various WPB Divisional Requirements Committees to direct product production and distribution and to allocate individual products under control were constantly impeded by these conditions. Adding and duplicating machines, for instance, as used in Service Equipment Division in early reports, included both electric and non-electric types. Sewing machines, on the other hand, were reported under the headings: "electric," "hand," and "treadle." Benzene as originally reported by the Chemicals Bureau meant both the pure pharmaceutical grade and the much lower motor grade. Hot water storage tanks defined by the Plumbing and Heating Division included four products ranging in weight of contained carbon steel from 50 to 1700 pounds, while track laying tractors, on the other hand, were subdivided into five weight classes and two models. Domestic electric mechanical refrigerators were reported by the Consumers Durable Goods Division in terms of units, regardless of size or weight. Such lack of system was widespread and was justified by administrators as an essential compromise of administration.

The need for better definitions of individual products coincided with the demand for establishing a usable classification to cover the entire spectrum of individual manufactured products. The problem of obtaining better definitions and classifications of individual products and the industries which manufactured them was not new to government statisticians. Government agencies such as the Bureau of the Census, the Bureau of Labor Statistics, and the Bureau of Foreign and Domestic Commerce had devoted considerable attention to the problems of properly classifying industrial data. Each governmental agency, however, maintained its own classification system and a given product or establishment was often classified so differently that data collected by one could not be related to that of the other.

To grapple with this problem in the prewar years, a committee was formed in 1934 under the sponsorship of the Central Statistical Board. The work of this group resulted in the publication in 1937 of the "Standard Industrial Classification."[1] When this volume was completed a "Standard Materials Classification" was begun but was not available in the early defense period when emergency requirements for standard materials classifications became important. By 1941, classification was a problem of considerable magnitude in the work of the OPM. In its efforts to create meaningful summaries of supply contract awards and the volume of preference rating certificates the "Standard Industrial Classification," as modified in the 1939 Biennial Census of Manufacturers, was the classification system used.

The groupings used in the Biennial Census of Manufacturers, however, did not provide for an adequate coverage of all the poducts involved in wartime production. As a result, the Division of Statistical Standards of the Bureau of the Budget (successor to the Central Statistical Board), with the aid of the WPB, The Bureau of the Census and other federal agencies represented on the Federal Committee on Economic Statistics, undertook a cooperative effort which resulted in the development of the "Defense Program Product Classification."

---

[1] At the end of the article.

This system subdivided many of the broad categories previously used and added new classifications for war products and their complementary items. Although the "Defense Program Product Classification" provided a foundation for the summarization of economic data, it was deemed necessary to group the data further to convey information to top executives. As a result, when the first PD-275 was issued in January 1942, the established classifications were grouped into 111 new product groups covering manufacture in the metal-working industries. In addition, seven new groups were established covering such activities as shipyards, naval air stations, and proving grounds.

Following discontinuance of form PD-275 this classification system was expanded and modified to tabulate form PD-25A, used under the first comprehensive metal allocation system adopted by the WPB. Under the product classification system used for PD-25A it was possible to classify production data with the specific product to which they related, to summarize data to reflect wartime industrial operations, and to provide summary information which the WPB could use for product administration.

Despite the great improvement in product classification for tabulation purposes, considerable difficulty arose in the WPB in attempting to use PD-25A data for administrative purposes. The core of the problem rested in the fact that the WPB internal assignment of product responsibility varied considerably from the product arrangements used in tabulation. In addition, WPB administrators felt that they needed more detail than it was possible to get from manufacturers' PD-25A schedules. Both problems could have been solved, but for these and other reasons this classification was seldom used for day-to-day administrative purposes following the discontinuance of the PRP and its form PD-25A in the spring of 1943. For tabulating such comprehensive general policy forms as WPB-732, CMP-7 and WPB-517, however, this system, continuously modified and improved, remained in use throughout the war period.

With the introduction of the Controlled Materials Plan in the spring of 1943 a new product classification had to be devised to fit its operational needs. Products were divided into two groups: Class A and Class B. Class B products originally were those in the component class but eventually included end-products as well. These were products the controlled material allotments for which were made by the WPB. All other products were Class A products and for the most part allotments were made to them through Claimant Agencies. Over and above the

Class A products were Claimant Agency programs, such as tanks, destroyers, and submarines. Both the Class A and Class B specifications of products were designed essentially for administrative purposes without regard to the logic of product nomenclature.

Material Nomenclature and Classification

The evolution of material definitions and classification paralleled that for products. Efforts to develop adequate information about basic materials brought to a head confusion which had existed as far back as the early NDAC days. There was little consistency of ideas about what constituted a basic material or what shape and form of elements composed it. Each industry had its own definitions. So did early war administrators.

Even a basic concept such as steel ingot was frequently misunderstood, misapplied and misused. Economists and statisticians, for example, in talking about steel thought fundamentally in terms of ingot and spoke continually of the capacity of American steel producers as being 60 million tons of ingot per year. In contrast, the steel industry men talked about finished steel, for example, plate, structural shapes, bars, sheets, etc., but knew that in their conversation tons of ingot meant 75 per cent or less of comparable tons of finished steel shapes and forms. Various operating divisions of the early war agencies were staffed with steel consumers who visualized tons of steel as finished products. They spoke of steel in material tons which translated ingot into finished fabricated component or end products. In this sense an ingot, depending upon the product, meant much less than 75 per cent of comparable basic ingot tons. Throughout the 1940–1941 period there was continual confusion in discussions of potential supply and demand for steel as a result of these diverse concepts.

The language of basic materials was by no means fixed in definitions pertaining to broad categories of basic materials such as the distinction between carbon steel and alloy steel, between alloy steel and stainless steel, and among other grades of steel, non-ferrous metals, and non-ferrous minerals. Thus, for example, alloy steel, as the steelmaker thought of it, started at the point where the carbon content was roughly less than .5 of 1 per cent and the manganese, silicon, copper or other alloy content was about 1.65 per cent. For many consumers, however, such low alloy figures were insufficient to meet their concepts of alloy steel. Stainless steel was distinguished from alloy steel by consumers but was generally considered alloy by steel makers. Actually, many people in the steel industry loosely defined as carbon steel a substantial part of the types of steel which non-technicians regarded as alloy. The reason was that low alloy content, particularly manganese, had always been recognized by steel producers as higher grades of carbon rather than alloy. Among steelmakers themselves the distinction between carbon and alloy steel was not always clear-cut because it depended on processing techniques as well as chemical content. In many cases, what was carbon to one mill was alloy to another. And peacetime concepts were modified to fit wartime specifications, further blurring the distinction in common usage.

Even more confusing was the nomenclature of basic material finished shapes and forms. Once again, difficulty sprang from varying popular usage in industry and a lack of familiarity with established nomenclature on the part of those whose experience in industry was limited. In the steel trade, for example, any finished mill shape more than 3/32 of an inch in thickness and 47 inches or more wide was considered a plate and anything under these dimensions was a sheet. To a manufacturer or a steel barrel or drum, however, all drum stock might be considered sheet whether or not it was over or under these limits. Some shipbuilders ordered "plates" when specifications carried a weight of over 25.5 lbs., per square foot. To some consumers, plates were all sizes over 7.65 lbs., per square foot and over 48″ wide; to others, sizes over 10.2 lbs., per square foot; and so on. On other steel shapes also, such as rod and bar, structural shapes, pipe and tube; in copper shapes; in aluminum shapes; and in other materials; usage was far from uniform. What was one man's sheet was another man's plate; what was one man's rod was another man's bar; what was one man's pipe was another man's tube.

These differences were reconciled under the early M Orders by requiring each producer or consumer of the material to report information in the terms specified on the form, irrespective of the manner in which records were kept. Beyond a particular allocation system, however, definitions and classification remained confused. A manufacturer who had to report comparable information to several product administrators found that each used different definitions and classifications. A producer of welding electrodes, for example, found that in reporting his consumption of alloying elements the definitions of tungsten, vanadium and molybdenum alloys changed with each person with whom he dealt.

Such problems not only created difficulties for manufacturers but raised obstacles at many points in the administrative process. Product administrators found it difficult to unify their treatment among different manufacturers using the same product, a situation which was reflected not only in their administration but also in top policy. In reports of administrators to top staff, separate product administrators often came to startlingly different conclusions. It was often difficult for policymakers to know whether one report was more accurate than another, whether one report used poorer figures than another, or whether the conclusions of various reports were really identical but expressed differently because of differences in terminology.

Complications increased as questionnaires for each material under allocation demanded greater detail. The tighter supply became in relation to demand the more detail administrators required. Under M-1-i, for example, aluminum producers were asked to report on Form PD-8 only their primary and secondary monthly shipments by nine preference rating categories. This form was promptly replaced by PD-26 upon which aluminum producers, refineries and smelters were asked to report on twenty specific aluminum shapes and forms. Through continuous expansion of reproting requirements aluminum producers were finally asked to report on CMP-24 a total of 43 shapes and forms. Eventually a listing of 77 aluminum shapes and forms was developed but the number for which reporting was required did not go beyond 43. Comparable expansion in other basic materials such as steel, nickel, copper, magnesium, cork and rubber also took place.

In September 1941, a movement began for the compilation of a standard materials list which would define precisely the meaning of basic materials, grades, shapes and forms and other species of the same genus. By this time the work of the Bureau of the Budget, noted above, had resulted in the publication of the *Standard Materials Classification*. A new materials list was developed from this volume and became known as *Materials List #1*. It embraced all of the important metal shapes and forms and included in varying detail such important commodities as textiles, lumber, miscellaneous minerals, and chemicals.

When this list was published the Army made known that it also was preparing a standard materials list. When the Army and the OPM lists were compared many inconsistencies were revealed. By means of compromise, modifications, and adjustments, the official "Red Book" or *Official Classification List of Raw and Basic Industrial Materials* was issued in February 1942 as a joint publication of the ANMB, the WPB, the Navy, and the Maritime Commission.

Crystallization of common usage by the publication of the "Red Book" solved the basic problems of material terminology. Thereafter the battle was more one of improvement than of struggle with basic concepts. Nevertheless, the use of material shape and form classifications on various reporting instruments varied widely. Items on *Materials List No. 1* were used for the first time on form PD-25A. Throughout the year 1942 and in the early part of 1943, PD-25A continued to incorporate this classification. In the meantime, various questionnaires used in connection with M Orders, inventory controls and P Orders continued their expansion of material shape and form reporting requirements. Since the use of neither *Materials List No. 1* or the "Red Book" classification was compulsory, material breakdowns on these questionnaires often varied considerably from the standard listings and PD-25A.

The movement towards an expanding list of material shapes and forms for copper, aluminum and steel allocation controls was cut short with the introduction of the CMP. For allotment purposes the CMP materials listing was short. The 85 shapes of carbon and alloy steel used on PD-25A were reduced to two CMP groups; carbon steel and alloy steel. The 24 copper and copper-base alloy shapes and forms of the PRP were reduced to four under the CMP, and the 21 aluminum shapes were lumped into 8 shapes and forms.

Other materials allocated by the PRP such as magnesium, zinc, nickel, and lead were not included under the CMP. But by the time the CMP was introduced their classification had become more or less standard. Condensation of shapes, and forms, such as took place under the CMP, did not occur with most other materials under individual allocation programs.

The decision to telescope basic material shapes and forms resulted from a combination of circumstances. To begin with, an expanded list of shapes and forms such as that used under the PRP would have made CMP completely unworkable, particularly in light of the first proposal to issue allotments on a monthly basis, to collect detailed bills of materials for all products and programs, and to ticket each allotment with a document as it filtered through industrial processes. Second, there was a natural revulsion from the lengthy listings included in PD-25A which became cumbersome both from the standpoint of manufacturer response and WPB administration. Third, experience with previous metal-allocation systems had demonstrated that the WPB's control, at least as far as authorizations were concerned, could be just as effective on the basis of broad material groupings as on that of detailed shape and form breakdowns. Individual controls could be used to resolve mill production and distribution problems for tight shapes and forms.

Not all the CMP operations were in terms of the telescoped groups noted above. In the original draft of the CMP there was established material shapes and forms under each main controlled material as follows: carbon steel, 13; alloy steel, 10; copper, 9; and aluminum, 33. Some CMP and related forms required reporting in variations of these shapes. CMP-26 (steel producers' daily reports of accepted authorized controlled material orders for carbon and alloy steel), for example, listed 46 carbon and 33 alloy steel shapes and forms for reporting purposes. CMP-22 (producers' report of shipments of steel castings) required data concerning 10 grades of castings. Aluminum suppliers reported shipments and unfilled orders on CMP-24 by 43 shapes and forms, and cumulative authorized orders accepted for delivery on CMP 12 by 22 shapes and forms. For copper the same sort of situation existed. On form WPB-3112, (brass mill application for copper, copper-base alloy, ingot and scrap) reporting was in terms of 10 different types and on WPB-3508 (brass mill program report) the respondents reported in terms of 4 brass mill shapes. These reporting instruments were typical of a large number of forms pertaining to mill shipments, applications and requirements.

There is considerable doubt whether telescoping of shapes and forms under the CMP was completely desirable. While it is true that summarization of shapes for the three controlled materials eliminated a substantial amount of posting, the absence of detailed shape and form information led to rules and regulations which might not have been necessary with more basic factual detail. Condensation under the CMP was partly based on the assumption that shape and form details would be available for products on prototype bills of materials. But few accurate bills of materials were ever received. No particular problems were created so long as fabricating facilities were balanced with the requirements of end programs and mill capacity was sufficient to satisfy the really essential demands. Under the operation of the CMP, however, there was a constant shifting of the degree of tightness among various material shapes and forms, in and out of detailed sizes, and in various shape qualities and characteristics. This experience necessitated the introduction of new methods to plug the various holes when they opened. The net gain, therefore, was not so great as might appear at first.

General Terminology

Precise meanings of many terms and phrases had to be clearly stated in each questionnaire to assure uniform responses. The problem of nomenclature, therefore, was more or less as broad as the complete lexicon of the spoken language.

WPB questionnaires showed continual improvement from loose and ill-conceived instructions on many early questionnaires to quite clear-cut and precise instructions on later questionnaires, oftentimes reinforced by reference to well-defined regulations and orders. But no matter how carefully instructions were written, they were always subject to criticism, misinterpretation and misunderstanding. On the one hand were those who objected to the "legalistic" and precise terminology used, and on the other hand were those who deplored the looseness of phraseology. Some manufacturers objected to the insufficient use of technical and "shop language" whereas others complained of phrasing which was subject to misinterpretation. Our interest here is centered largely in a recognition of such problems together with a few typical illustrations of points of difficulty, rather than in a definitive analysis of either the problem or of its solution.

The term "capacity" was continually misunderstood on WPB reports. On WPB-3002 (manufacturers monthly report of unfilled orders and schedule of shipments) and WPB-2578 (report on speed reduction units), for example, manufacturers were required to show in units their productive capacity and schedule of deliveries for each type and size of product. Producers of electrical testing equipment, resistors, electric motors and generators, speed reduction units and unmounted gears inveighed against the term "capacity." They insisted that since two or more models could be manufactured with the same facilities and the same wage earners, the productive capacity for any given size and type depended upon the facilities in use at any given time, and that one month's operations could require heavier production of one model and lighter production of others, while in the next month the situation could be reversed. Even though instructions often stated that capacity by types should follow the proportions established by past actual production and that all capacity data ought to be concurrent, some manufacturers maintained that productive capacity for each model was something which did not exist in fact. Other manufacturers alleged that the term "capacity" was too indefinite. They were not certain whether it meant theoretical capacity, practical operating capacity under the existing conditions of material, labor and machine tool utilization, or actual capacity based on past performance. Although the definition of the term was apparently clear to those in the WPB responsible for its use (the term generally meant the maximum monthly rate of production of each item which could be reached and maintained with the equipment available, assuming the availability of materials and the maintenance of maximum hours of operation) to industry the term lacked meaning.

The phrase "delivery schedule" raised considerable confusion. WPB-1801 (production of electric motors and generators), for example, requested for each model whose production was reported the standard delivery time in weeks. Industry replied that there was no such thing as a "standard delivery time" for a particular product. To have significance to a manufacturer, the phrase "delivery schedule" had to relate to shipments, since a manufacturer scheduled his production against dates of shipment from his plant and not dates when delivery was made to a customer. Closely connected with these words was the term "requested for delivery." WPB-3002 and WPB-2578 (noted above), for example, called for units of product or dollar value scheculed for shipment and requested for delivery by months. Instructions asked the respondents to specify the total dollar value or number of units included in the total unfilled orders broken down by the amounts that purchasers requested in each of the indicated time periods. Industry insisted, however, that it could not develop the data from existing records, nor could it see how the data could be developed. This resulted from industry's interpretation of the data request to mean the amounts originally requested by customers before placing their orders, information which of course was unobtainable. The intention of the WPB was to compare the items scheduled for shipment with those requested for delivery so as to be in a position to determine the extent to which producers were falling behind schedule.

These illustrations are no more than indicative of a problem which baffled both the WPB and industry throughout the war period. There were hundreds of other words and phrases which caused considerable difficulty, such as idle, excessive, obsolete, usable and nonusable inventory; net shipments; principal product; captive operations; maintenance, repair, and operating supplies; repair parts; and so on. Experience demonstrated that a considerable period of time was often required before words and phrases were so defined that they were mutually comprehensible by the WPB and industry. It also left little doubt that for best results terminology used in questionnaires had to conform as closely as possible to that generally used in industry and that the more precisely a term or phrase was defined the less opportunity existed for misinterpretation and consequent inconsistency in data reported by manufacturers and used by administrators. But what conformed to industrial usage and what was precise were determined only by long and hard experience.

Industrial Reporting Units

An important, but often overlooked question was that of the precise area to be covered by a single questionnaire. One of the basic common denominators determining the usefulness of collected information was the industrial unit for which reporting was to be made. The basic common denominators in practically every industrial statistical report of the WPB were: (1) individual products (or materials), (2) individual industrial plants, and (3) companies. The structure of industry, and in turn the subject matter of production and distribution control, demanded that this be so.

The fundamental distinction between a product (or material) report and a plant report was simple. All questions in a product or material report related to specific products or materials; individual products or materials constituted the subject of the report. The plant report on the other hand used an industrial establishment in one geographical location, or a subdivision of the plant, as the broad subject and required information about all products, all materials and other productive elements receiving attention within the plant area. As such, of course, a plant report was an aggregation of product or material reports but covered all products and materials as well as other relevant productive elements of a plant. In some plant reports all reported data about a plant related to individual products made by the plant or individual materials used in the plant. A company report was often a consolidation of either individual product or plant information, seldom the latter, on a company-wide basis irrespective of geographical location of individual plants or the production locale of individual products or materials.

These distinctions between plant and product reports were hazy in the earliest surveys. Fundamental problems confronting the OPM in early 1941 concerned industrial capacity for certain products or materials, inventories, and production and use of a limited number of individual materials and products in short supply. Early questionnaires sought facts about individual products and materials without a conscious distinction between the plant and the product approaches to informational requirements. Reports had to be sent to plants or companies for completion, not to products. If the needed information related to capacity or to inventories the report generally used a plant as the reporting unit, for data relating to such subjects were reportable and had meaning only in terms of a plant. If, on the other hand, the schedule concerned individual products or materials the result usually was a product report which confined questions to an individual product or material. Frequently, however, these reports (OPM-6 and OPM-23), for instance, combined questions as to capacity with data requests for individual products or materials with the result that the distinction between a product and a plant became blurred and a questionnaire often in fact used both a product and a plant as a basic reporting measure. But as the war progressed the distinction between the two reporting bases became clearer.

Reporting by industry in these fundamental units was seldom uniform from survey to survey for the same questionnaire or different questionnaires. Thus, for example, specific data on one plant report, such as WPB-732, might include a universe slightly different from that reported on another plant report, such as PD-25A. Great variation existed from report to report because of flexibility permitted in reporting and variation in the field of activity covered by a given set of industry records. Even greater variations existed in factual reporting for specific products.

Product and Material Units of Measure

Closely allied to the names and classifications of products and materials were their units of measure. Here again the WPB found widespread differences in reporting of standard quantities and values. And here again the trend was towards greater precision in definition; and more general use of those standards which it was found were mutually understood by both industry and government, readily reportable by industry and acceptable for administrative purposes. Measurement units generally were necessarily fixed according to industry usage and industry records, particularly when they related to past activities. Unfortunately, however, usage was not always standard among plants and industries and usage common in industry was not always of most utility to government.

Material units of measure generally fell into two distinct groups, or else defied classification. The most widely used was a weight measure. Most basic materials, particularly the metals, originally requested reporting in terms of tons or pounds. Even these simple concepts created difficulties. Some questionnaires requested information in long tons, others in net tons, and still others in short tons, generally without defining the standard specifications pertaining to the concept used. Some forms mixed the reporting requirements, such as PD-209A (report of vanadium consumption), and data were requested in both long and short tons with no definitions of either. The result was a 12 per cent variation in reporting. Experience led the WPB to request reporting for most materials, when tonnage was specified, in short tons of 2000 pounds each. Reporting in pounds presented no particular problem because a pound was solidly entrenched as a standard unit.

For those materials which were incorporated in other basic materials at the mill level, particularly the ferroalloys, reporting was often required in terms of the content of these other materials. PD-358 (producers application for molybdenum for melting), for example, requested the per cent of molybdenum in finished products. PD-359 (producers report of molybdenum containing metal) requested reporting by pounds of molybdenum content. Complications developed from time to time in the use of such measures. Nitrogen-containing materials were at first measured in tons of ammonia content, then ammonia equivalent, and finally in tons of calculated nitrogen content.

Units of measure outside these areas were too diverse for ready classification. PD-338 (inventory of jewel bearing material) measured jewel bearing materials, such as natural sapphire and synthetic ruby, in carats. PD-236 (consumption of jewel bearings) established reporting in terms of the number of jewels contained in bearings. Mercury was measured in flasks, phosphates in terms of equivalent phosphorus pentoxide, ethyl alcohol in gallons of 190 proof, penicillin in billions of oxford units, and leather in terms of equivalent hides and skins.

Three main types of units of measure were commonly used in connection with products. The first and most desirable unit for products, and the one most frequently used, was one unit of a multiple of the product in question. This was satisfactory when the product was uniform, such as hot drink cups, telephones or kerosene railroad lanterns, but could not well be used in the case of products presenting diversity in important characteristics. The unit classification was strained in numerous instances in summary tabulations, as in the case of steam and diesel locomotives, scales and balances and various machines. Second, many orders established controls on the basis of the tonnage of iron and steel or other metal used in the production of the controlled products during a base period and reports were made in terms of units, pounds, or value of the finished product. This material content approach, of course, was limited to availability of records and was not as widely used as the other methods. Third, when the use of units and size or weight measures was unsatisfactory, dollar value measurement was used. This was inherently unsatisfactory as a control measure because of lack of information as to quantity and quality of production and distribution, number of units, quantity of raw material and labor required, and changes in prices. It was used in lieu of anything better and at the same time was practical.

Confounding problems often arose because of the use of varying units of measure and the difficulties of settling upon a satisfactory conversion formula to secure comparability of related data. The Foreign Economic Administration, for example, received from the WPB allocations of leather and textiles in units of equivalent hides and linear and square yards and pounds of cloth or yarn, but issued licenses and made requisitions in terms of shoes and finished textiles. Months were required to develop conversion factors and standardize terminology and measurement units. Comparable problems were encountered in many other products and materials.

Difficulties with varying units of measure resulted in a trend in the WPB, continuing through the war years, to require reporting of material and products in terms of individual units such as quantity units for products and tons and pounds for metals. The coincident development of product and material definitions and classifications when related to uniform units of measure made the data collected by the WPB increasingly more useful with each step forward in these areas.

Time Periods Covered on Reports

Under the rapidly changing conditions of a war economy a year was too long a period to wait for reports upon which to base policy, and time periods moved from quarters, to months, to weeks, and often to days depending upon the urgency of a problem. From time to time products were allocated on a daily telephonic basis. Broad material allocation programs, however, could not operate on so short a time period. Many of the individual allocation systems in metals existing up to and through the Production Requirements Plan operated on a monthly allocation basis. The PRP itself was based on quarterly reporting and allocation. This, it was felt was too long for many tight items, and the early promulgation of the CMP envisioned monthly allocation. It was quickly perceived, however, that such a time period would involve a tremendous amount of paper work, conferences and decision writing. As a consequence the time period instituted with the first operation of the plan was the calendar quarter. As circumstances permitted, other allocation plans quickly followed the quarterly allotment procedure, and the calendar quarter finally evolved as the best time unit for making allocation, scheduling production and reporting on implementation.

Measures and Definitions of End-Use Data

The general approaches to production control embodied in the preference rating system, P Orders, and M Orders, encountered a problem relating to identification of end programs at sub-manufacturing levels which frustrated the WPB throughout the war period. Early operators of material allocation controls recognized that their decisions concerning mill shipment authorizations were not particularly intelligent in the absence of information relating to the uses, particularly the end-uses, to which the basic material ultimately was to be put. This type of information was difficult to acquire because no workable system existed for passing information from the top to the bottom of the industrial system.

There were many different approaches to the development of end-use information. Throughout the year 1941 and into early 1942 constant difficulties with old methods resulted in their revision or the institution of completely new systems, each striving to obtain information about end uses which was considered vital to the operation of each allocation plan. By the spring of 1942, after almost one full year of experience, methods to derive end-use information were little closer to their goal than in the spring of 1941.

The approaches current in the spring of 1942 may be classified into several rough categories. (1) The *general* approach was adopted in PD-69 (consumers pig iron order for shipment) which requested applicants to give a complete description of the ultimate use or product to be made from pig iron for which application was made. With little help, applicants were asked to be very specific, i.e., "Pump castings for Navy ships." (2) PD-123 (producers report on copper and copper alloy deliveries) embodied what may be classified as a *use* approach. This form originally contained 37, and ultimately 100, classifications of end uses for which producers were to report deliveries of various copper and copper-alloy shapes and forms. On the basis of the classification given it was not only impossible for producers to file completely, but one end product conceivably could appear in three or more classifications as set forth, depending upon its ultimate use. As a result, tabulations of this form were often weird and meaningless. (3) The *Product-type* approach may be illustrated by PD-180 (customer application for phenol). The applicant was obliged to enter the name of the product into which he was incorporating phenol and the specific ultimate use of that product. (4) The *use and product* approach was typified by PD-197 (producers report of lead and lead alloy shipments). On this form 33 classes of products were preprinted in the stub, such as storage batteries, cable covering, and white lead. A cross-classification was requested showing the amounts of lead used in each of the preprinted products.

The persistence with which this type of data was sought and the willingness with which the WPB introduced rapid changes in past methods, each more complicated than its antecedent, testify to the importance attached to this sort of information and the desperation with which it was sought. Neither the material procurement agencies nor the WPB commodity divisions were in a position to match demand-supply relationships on a factual basis. As a result it not only was impossible to assign the resources of the country to those uses of most importance, but it was difficult to determine the desirability of expanding production of critical materials. The key to the puzzle seemed to lie in the development of machinery by which accurate identification of end products manufactured by prime contractors could be transmitted throughout the entire industrial system, through intermediate production and components, to the basic metal mill producers. Such a system, it was reasoned, would allow the WPB to follow the distribution of basic materials and components of each important end-product and assure its completion on a balanced schedule and in conformance with a carefully pre-arranged principle.

To meet these objectives a bold experiment in end-use identification was prepared in June 1942 with the issuance of Priority Regulation No. 10. The regulation observed that its purpose was "(1) to obtain standardization and reduce the rules of different forms for allocation purposes that industry now must submit to the War Production Board, and (2) to furnish information needed by the War Production Board in the allocation of materials. . . . To allocate intelligently," it went on, "it is necessary to know the subdivision of the war, industrial and civilian programs for which materials are going to be used. For example, it is essential to know whether the materials are going to tanks, or destroyers, or railroads, or to office machinery and supplies in order to determine allocation policy. In addition, it is necessary to know in a general way what type of purchasers will eventually receive particular products of industry, i.e., whether the product will eventually be delivered to the Army, Navy, etc."

The theory behind the code was simple. Each purchase order from a prime contractor would include an end-use code indicating the use to which the purchased materials or products were to be put. Each subcontractor would then repeat the code symbol on his orders for materials and parts until, in theory, the symbol appeared on orders placed for basic materials. Suppliers of critical materials would then be in a position to report their shipments by end use codes, which would reveal information on both the end use and end user. It was recognized that this plan would not be feasible in every case, but it was thought that the band of unidentified end use could be narrowed and that the system would provide the first real check on the relationship between end products and scarce materials.

The WPB materials and commodity Divisions violently protested the use of this system. Having developed end-use reporting requirements to fit their particular needs, they insisted that the classifications in the end use code were so broad as to be almost meaningless. The purpose of this code was to standardize the many existing codes and to keep the number of classifications to a minimum, so as to reduce the burden on industry. Yet if the code did not meet the needs of the Divisions it was doomed to failure. To say this is not to say that the end use codes developed by the Divisions were workable. On the contrary they lacked uniformity and could not be tabulated. But they did provide the narrow bands of end use which the Divisions wanted and which the newly devised code lacked. The theory behind the code may have been sound, but it fell down in its practical applicability.

Serious criticism also came from many industries. The steel industry, for example, complained that it was faced with serious difficulties in the application of the regulation to its normal business practices—so much so that the data received by the WPB from steel mills had little value for allocation purposes. Fundamental problems of the steel industry lay in three areas. First, steel warehouses were required to analyze monthly all their purchase orders to determine the allocation classification pattern which they would use on their own purchase orders which they placed with steel suppliers in the following month. Since the warehouses received thousands of small orders each month, the accounting and tabulating problem was staggering. Moreover, since the Steel Division of the WPB thought it could not and therefore would not allocate steel in accordance with the information furnished by the classification system, all the paper work involved was useless. Second, since the warehouses added extra charges for cutting, handling, etc., analysis of the sales pattern by dollar volume as required by the regulation, was misleading. Third, since the Steel Division of the WPB required information according to the nine purchaser classifications used under its own Order M-21 instead of the five purchaser symbols used under PR-10, the industry had to keep two sets of records.

Following a short but turbulent life the system was abandoned in November 1942 as a failure. It was cumbersome in use, unworkable in large segments of industry, particularly at the component level, and produced unreliable order identification on mill order books. The concept was valuable but the methodology was poor.

Throughout the experience with Priorities Regulation No. 10, it was becoming clear that satisfactory control over war production could not be answered by a suitable end-use system slone. It was but one ingredient in a far broader need which embraced the balancing of end programs among themselves in relationship to the totality of scarce materials; the scheduling of materials through the industrial system to end products, in terms of quantity and time; the adjudication of competing claims at critical levels of fabrication and direction of mill fabricating facilities.

Following the failure of Priorities Regulation No. 10, the Production Requirements Plan provided part of this needed method. The introduction of the Controlled Materials Plan in 1943 furnished a practical mechanism by which end program symbols could be transmitted down through the contractual chain to mill levels for controlled materials. The plan accomplished this simply by requiring that a controlled material allotment and program symbol had to accompany each purchase order if the product ordered used any one of the controlled material shapes and forms. Not only were prime consumers requested to use major program symbols to identify orders carrying controlled material allotments but they also identified rated orders for class B products for which they did not reallot quantities of controlled materials. By means of this procedure end-program identification, based on CMP allotments, was provided for a very large part of American industry, from procurement agencies through to metal mills.

The system by which Claimant Agency symbols were passed down to Class B producers was not designed to provide precise identification of either Claimant Agencies or programs. To begin with, many manufacturers operating under P, U, and other orders and regulations used a variety of specialized symbols not identifiable in terms of Claimant Agencies. Second, rated orders from prime consumers did not always carry Claiment Agency symbols. Finally, extension to Class B producers by prime consumers was in terms of Claimant Agencies themselves, and books were not maintained showing program symbols. The result was a broad specification of end-use on CMP-4B applications. Despite many difficulties in identification growing out of weaknesses in parts of the system, the Controlled Materials Plan solved practically, from an administrative point of view, the end-use problem for controlled materials. The solution, however, was not so much imbedded in a method of classification as in a complete new comprehensive allocation system.

Individual allocation systems continued to exist for materials and products other than those under general comprehensive programs such as CMP, Lumber, and Rubber Tires. The methods for determining end uses were compromises between the desires of the WPB and the records of industry. As might be expected the methods varied considerably and incorporated the same features which were used under M Orders. As a consequence they engendered the same pre-CMP problems of compliance, the same type of industry criticism and the same collection of inadequate data. For many materials and products the problem was never solved satisfactorily. The discipline created by Priorities Regulation No. 10 and the operation of the CMP helped, but industrial practices negated passage of complete end-use information to raw material sources without a superimposed method such as the CMP allotment chain.

The length of time required by the WPB to develop satisfactory definitions and classifications should not be viewed too critically. Of course the struggle was much influenced by the fumbling which characterized the entire evolution of production and distribution controls. But the matted problems of definitions and classifications were, even beyond that, hard to disentangle. Even when definitions were clear cut new circumstances inevitably arose to blur them. In some fundamental aspects, the problem of nomenclature specification was insoluble. What was a sound definition depended upon the persons using it, industry usage, the state of classification systems and the precise needs of a given control policy in which it was a part. The broad definition of a material such as carbon steel, to cite one instance, was reasonably satisfactory in the early war period. But it really meant little. To speak of carbon steel sheet helped a little, but not much. Further clarification of carbon steel sheet in terms of a specific gauge, size and shape was better, but still not sufficient for complete clarity. To carry the needed extensions of definition to extremes, which would be necessary for complete and unequivocable understanding, obviously would have entailed the most minute description of characteristics of a product or material. Even this would not be enough. In a war period, rapid changes take place in old products and there is a constant introduction of new products with new names, characteristics and uses. In such circumstances workable definitions and classifications are all that can be expected. Much time and experience, however, are required before this condition can be reached.

[1] *Standard Industrial Classification*, Central Statistical Board (through 1937); Bureau of the Budget (1940 to present). Part I, Manufacturing Industries, December 10, 1937. (Tentative draft.)

Volume I, Part I, Manufacturing Industries, List of Industries, May 10, 1938.

Volume I, Part I, Manufacturing Industries, List of Industries, 1939.

Volume I, Part III, Manufacturing Industries, Alphabetic Index of Products, Establishments and Processes, 1939.

Volume I, Part IV, Manufacturing Industries, Alphabetic Index of Products, Establishments and Processes by Major Industry Groups, 1939.

Volume II, Part I, Non-manufacturing Industries, List of Industries, 1939.

Volume II, Part II, Non-manufacturing Industries, Description of Industries, 1940.

Volume II, Part III, Non-manufacturing Industries, Alphabetic Index of Products, Establishments and Services, 1940.

Volume II, Part IV, Non-manufacturing Industries, Alphabetic Index of Products, Establishments and Services by Major Industry Groups, 1940.

Volume I, Part II, Manufacturing Industries, Description of Industries, 1940.

Volume I, Manufacturing Industries, 1941.

Volume II, Non-manufacturing Industries, 1942 (still current).

Volume I, Part I, Manufacturing Industries, Titles and Descriptions of Industries, November 1945.

Volume I, Part II, Manufacturing Industries, Alphabetic Index, December 1945.

Volume II, Non-manufacturing Industries, 1942, available at GPO for 25¢.

Volume I, Part I, November 1945, available at GPO at 75¢.

Volume I, Part II, December 1945, available at GPO at $1.00.

Standard Classified List of Commodities, Prepared under auspices Executive Office of the President, Bureau of the Budget; War Production Board; and Department of the Treasury, Division of Procurement. First edition (for comment only), July 1942. Reg. First edition, December 1942. Revised edition (Vol. 1), May 1943. Vol. II of Alphabetic Index, February 1944. Supplement to Vol. II, November 1945. July and December 1942 reports not available except in the files of the Bureau of the Budget. Other reports available at U. S. Government Printing Office (May and Feb. 1943 reports @ $1.25, and November report @ $2.00.)

BOOK REVIEWS

Edited by

OSCAR KRISEN BUROS

Rutgers University

The Teaching of Statistics: A Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics. *Harold Hotelling,* Chairman (Professor of Mathematical Statistics and Head of the Department; Associate Director, Institute of Statistics; University of North Carolina, Chapel Hill, N. C.), *Walter Bartky* (Dean of the Division of the Physical Sciences and Professor of Applied Mathematics, University of Chicago, Chicago, Ill.), *W. Edwards Deming* (Adviser in Sampling, Bureau of the Budget, Washington, D. C.), *Milton Friedman* (Associate Professor of Economics, University of Chicago, Chicago, Ill.), and *Paul Hoel* (Associate Professor of Mathematics, University of California, Los Angeles, Calif.). Reprinted from *Ann Math Stat* 19(1):95–115 Je '47. Washington 25, D. C.: National Research Council (2101 Constitution Ave.), 1948. Pp. 95–115. Paper. $0.25.

The Teaching of Statistics in Universities and Colleges. Based upon a report by the Royal Statistical Society's Committee on the teaching of statistics: *E. S. Pearson* (Chairman), *R. G. D. Allen, H. Campion, William Elderton, C. Oswald George, R. F. George, M. Greenwood, D. Heron, J. O. Irwin, H. Leak, E.C. Rhodes, R. Stone, L. H. C. Tippett,* and *J. Wishart.* Preprint of *J Royal Stat Soc* 110(1):51–7 pt 1 '47. London W.C.2: Royal Statistical Society (4 Portugal St.), 1947. Pp. 7. Paper. 6*d.*

REVIEW BY ALAN E. TRELOAR
Professor of Biostatistics, University of Minnesota
Minneapolis 14, Minnesota

SHORTAGE of qualified personnel within the statistical profession has been a matter of concern to employers demanding services from that source. The reports reviewed herein clearly testify that those charged with leadership in recruiting and training such personnel not only share that concern, but are acutely aware of grave problems basic to the unbalance which, if not removed, will lead to continued deterioration as well as wastage of educational opportunity in a very broad sense.

Appreciation of the usefulness of statistical methods was limited to a few farsighted pioneers at the beginning of this century. Response to their demonstrations of new power in analysis was very slow in developing. Biologists and workers in the social sciences generally were not able to assimilate the new biometry being created for them; for the most part they not only lacked the mathematical training so necessary to appreciation of its approach, but were most reluctant to believe that there was any hope that the variations which beset them were amenable to mathematical treatment. Only a few of broader vision saw great hope in the new tools and applied themselves to work on the frontier with tremendous enthusiasm despite their poverty in mathematics. With indefatigable energy they set an example in both teaching and research that gradually leavened the main lump. Their teaching was eminently practical. It had to be, for mathematical abstraction quickly bereft them of followers. Students learned by analogy how to apply the methods rather than to comprehend fully the basic principles of a new philosophy and the limitations imposed by the simplifying assumptions made by the theorists to reach the first solutions. These students have in turn become teachers, carrying responsibilities for which they were equipped only in a very limited sense.

In reviewing the present situation it is well not to lose sight of the contributions made by these pioneers with all their handicaps. Statistics enjoys a large demand today and is what it is largely because of them. Theoretical statisticians who now strive to serve the sciences and assume leadership responsibilities for the future development of statistics should properly temper criticism, no matter how justified it may seem from specific points of view.

The reports being reviewed herein, differing sharply in their emphasis on criticism, are in close agreement when considering what should be done in the future. Both reports were preceded by another* from the Committee on Applied Mathematical Statistics (L. P. Eisenhart, S. S. Wilks, C. I. Bliss, E. U. Condon, H. O. Gulliksen, L. J. Reed, C. F. Roos, W. A. Shewhart, H. M. Smallwood, F. F. Stephan) of the National Research Council, which urged the same objectives. All three should be considered together. The American point of view is thus more adequately represented, both in numbers and in balance of practical and theoretical components. The Royal Statistical Society report is based on the deliberations of many persons representing a wide coverage of both theoretical and applied statistics.

One need not review here evidence submitted of interest in and need for statistical methodology. A brief summary of the proposals for meeting the need would be in order, however. An introductory course of an orientation type is recommended for all college students as an essential component in a truly liberal education. Probability has replaced certainty in the modern way of looking at things, and those not familiar with the concept of aggregates and the fundamentals of the statistical argument in relation to it face a difficult mental reorientation in their efforts to keep abreast of life's affairs. The committees look forward to the day when this orientation will be a regular part of the secondary school education, but for the immediate future provision must be made for it to be given in the college curriculum. It is judged that the course would most profitably be built around experimental demonstration; it should not require special mathematical ability.

A basic course in statistical reasoning and methodology covering principles common to all applications and given centrally for all students needing more than just a general orientation is strongly urged. The present system of many departmental courses of all qualities is judged to be badly in need of overhaul; applications departments should yield the common ground, concentrating on facets peculiar to their own special fields, thus fostering interest and competence. It is acknowledged that the general course would have to be given for some time with two levels of mathematical prerequisites, but these would ultimately fuse with calculus as a prerequisite for all. The two courses given meantime should correspond in every way possible and have the same coverage.

---

* *Personnel and Training Problems Created by the Recent Growth of Applied Statistics in the United States.* Reprint and Circular Series Number 128. Washington, D. C.: National Research Council, May 1947. Pp. v, 17. Paper. $0.25. Reviewed in this JOURNAL (Dec. 1947), by George W. Snedecor and Holbrook Working.

Departments of application should continue to give statistical courses reconstituted to accord with the plan stated in the preceding paragraph. These courses might be regarded as intermediate in grade. They should embrace discussions of sources of data, limitations peculiar to them, and applications of methods with special emphasis on practical considerations. Attention would be given here to procedures of interest to the field of application but not of sufficient general importance to be included in the central introductory course.

Offerings at the advanced level would be planned in accord with facilities made available to meet existing needs. The IMS report attaches great importance to the training of future teachers through the route of advanced courses in statistical theory demanding advanced mathematics as a prerequisite. While all reports emphasize the importance of experience in application as an essential concomittant in statistical training at all levels, there is a note of warning in one (IMS) against allowing the specialist in statistical method to participate too extensively in applications lest they "so engross his attention that he cannot keep up with the progress of statistical method itself." History seems to this reviewer to indicate a drift toward theory by those of considerable competency in both theory and applications.

Centralized teaching of statistical method within any college, as envisioned by these committees, calls for creation of a special teaching unit whose personnel would be free to study statistics for its own sake. "In statistics, as in other subjects, a profession of general practitioners requires the support of an inner circle of specialists, research workers and teachers; and the inner circle maintains its strength partly through teaching and serving the general profession. . . . Where this advanced teaching can be associated with an important research school of statistics, considerable advantage will follow" (RSS). An independent department or institute of statistics is favored by all three reports, but where attachment to an existing department is insisted upon administratively the RSS committee recommends affiliation with mathematics. It is preferred that this unit have its own budget and be equipped as completely as possible with all facilities for practical work and research, including easy access to a fully adequate library of original sources and mathematical tables. This unit and the statistical units within applications departments should be drawn into close association through a coordinating committee, each gaining strength through reciprocal aid. Such a committee drawn solely from applications departments must not be accepted as a substitute for an administratively independent unit devoted primarily to teaching and research in statistical methodology.

The RSS report makes no comment on current standards of teaching in statistics. The NRC committee acknowledges as undesirable the wide variation in quality of statistics courses offered by applications departments within the colleges of this country, and the vested interests which the departments have developed in those courses, but goes no further. Only the IMS Committee has the courage (if courage it be) to attack forthrightly the procedures of appointment and administration of statistical personnel within the applications departments. An olive branch is offered in statements of the form: "All this does not mean that any considerable number of people teaching statistics are unworthy or objectionable members of the academic community. Many, indeed, are of superior intellect, upright character, personal charm, and undoubted teaching ability. Some are making creative contributions to other subjects. The only trouble is that they are teaching a subject in which they are not specialists, and which progresses so fast that only specialists can keep up with it." The dove of peace, laboring somewhat in its flight, makes no call at the offices of the responsible administrators. The argument is plausible, skillfully phrased, and reading it provides some entertainment. But this reviewer doubts that it will add to attainment of the objectives with which the report is so much concerned. A supply of adequately trained personnel in the sense defined has not been available. Theoreticians were very slow in responding to the need they are now so anxious to fill. They will be welcomed with unquestioning enthusiasm just as soon as they demonstrate that they can do the job more effectively than their predecessors. The affinity of theoretical statistics and mathematics is unquestioned. The report admits that mathematicians have so behaved in their teaching as to make children "resolve to quit mathematics as soon as possible." Nothing surely would be more welcome than clear demonstration that such performances will have no place in the techniques adopted by theoretical statisticians in undertaking the heavy responsibility of teaching statistics on a college wide scale.

To say these reports are important is to belabor the obvious. They constitute a most commendable attempt to plan for a better future. All teachers of statistics should read them carefully and reflect constructively on the problems they attack.

Investment, Location, and Size of Plant: A Realistic Inquiry Into the Structure of British and American Industries. *P. Sargant Florence* (Professor of Commerce); assisted by *W. Baldamus* (University Research Fellow). (University of Birmingham, Birmingham, England.) National Institute of Economic and Social Research, Economic and Social Studies VII. London N.W.1: Cambridge University Press (Bentley House, 200 Euston Road), 1948. Pp. xiii, 211. 18*s*. (New York 11: Macmillan Co. [60 Fifth Ave.] $5.00.)

REVIEW BY EDGAR M. HOOVER
Staff Economist, Council of Economic Advisers
Executive Office of the President, Washington 25, D. C.

PROFESSOR Florence reports here on an analysis of three interrelated attributes of industries: prevailing size of establishment, degree of geographic concentration, and invested capital per man. The basic data are primarily British and American production censuses of the latest available prewar years. He develops measures for each of the three attributes, seeks to explain the observed associations among them, and evaluates the significance of the findings for purposes of public policy.

The book as a whole represents a notable advancement of the state of our knowledge of industrial structure. Its value is enhanced by the presentation of industry-by-industry tabulations of nearly all the statistics discussed.

At times readability and directness appear sacrificed to wealth of suggestive detail. For example, some pages of text and one whole appendix are devoted to measures of the "heaviness" of industries, though the concept of heaviness is not used anywhere else so far as this reviewer could discover. On the whole, the reader will be grateful for the wealth of suggestive loose ends tagged for further exploration.

Size of plant is measured (after a discussion of alternatives) by number of employees. The extreme skewness of the size distribution of plants in an industry is circumvented by using frequency tabulations in which each class limit is 2 or 2½ times the next lower one. All the author claims in behalf of this inelegant procedure is that it seems to work: a distinct modal or "prevalent" size class appears in most industries except where wide variety of sizes is to be expected on a priori grounds.

The aspect of location with which this book deals is the deviation of the geographic distribution of a given industry from that of industry as a whole. The "coefficient of localization" has already gained fairly general acceptance as a result of its previous use by Florence and others, but it is here tabulated for the first time for a nearly complete list of industries of both countries.

The "intensity of investment" of an industry poses the most serious problems of measurement. Horsepower capacity per worker is chosen as the least objectionable of several possible indicators.

When specific British industries are compared with their American counterparts, a striking degree of international correspondence appears in plant size, location pattern, and use of power. This leads the author to his basic theses: that the basic technical operations of industries (and not merely the national political, social and financial milieu) are significant in accounting for interindustry differences, and that size, location, and use of power are interrelated characteristics.

In analyzing the relation between an industry's typical size of plant and its coefficient of localization, Florence uses simple contingency tables. The observed relation of plant size to localization is however, as earlier studies by Florence had indicated, a complex one. It goes considerably beyond the range of statistical proof as yet.

Plants in many of the most scattered or "ubiquitous" industries show no definite central tendency in size. In areas of dense concentrated market or material supply, plants can be larger than elsewhere, so a diversity of plant sizes is naturally found within a single industry.

In industries with intermediate degrees of localization, greater localization is associated with larger plant size.

The most highly localized industries are characterized by medium-sized plants because (*a*) in these industries transport costs do not limit size of plants, so plants need not be small to be close to scattered suppliers or markets; and (*b*) close juxtaposition of plants makes it possible to secure many of the economies of large-scale production by virtue of the size of the production center rather than of the individual plant—plants therefore can specialize more narrowly, and do not have to be large to be efficient. This hypothesis as to highly localized industries is tested by a more detailed analysis of concentrations of metal trades in Birmingham and textiles in Lancashire.

In exploring the relationship between size of plant and horsepower capacity per worker, Florence uses the same technique of contingency tables. In general, larger plants are associated with greater use of power per man, greater financial requirements, and greater prevalence of multi-plant firms.

The last part of the book speculates on implications of a continued trend toward larger plants and larger firms. There is little novelty in the author's case that protection of the interests of labor and the consumer will require a better and broader science of industrial management along with watchful public scrutiny or control of monopolistic industries.The most interesting point in the discussion of policy problems is an explanation of the difficulty of economically encouraging small business and dispersion of industry to small nonsuburban communities. According to the relationships shown by Florence's statistical analysis, most industries that are not tied by high transport costs to local markets or sources of materials are either highly clustered or characterized by large plants. Thus the promotion of small decentralized business faces the problem of increasing the efficiency of small plants or breaking up or transplanting complex localized groupings of plants. There may be further conflict between the aims of decentralization in nonsuburban small communities and local diversification of production.

These barriers are not necessarily impassible. The observed associations of size, location, and investment on which Florence bases his case are sufficiently loose to provide many exceptional industries and plants, and the economic advantages of size and localization are not immutable. Nonetheless the basic difficulties remain and this book has done a useful service by bringing them to light.

Foundations of Economic Analysis. *Paul Anthony Samuelson* (Professor of Economics, Massachusetts Institute of Technology, Cambridge, Mass.). Harvard Economic Studies, Vol. 80. Cambridge 38, Mass.: Harvard University Press (38 Quincy Street), 1947. Pp. xii, 447. $7.50. (London E.C.4: Oxford University Press [Amen House, Warwick Square], 1948. 42*s.*)

Review by George J. Stigler
Professor of Economics, Faculty of Political Science
Columbia University, New York 27

Professor Samuelson's distinguished doctoral dissertation was largely completed about a decade before its publication (p. vii). This delay in appearance is doubly unfortunate: unfortunate for other economists, deprived of portions of its brilliant theoretical analysis and criticism not hitherto published; unfortunate for Samuelson, much of whose subject matter is currently of much diminished interest. But the problems he deals with have been and will continue to be important and his treatment of them should be read by all mathematically trained economists interested in formal theory.

The leitmotiv of Samuelson's treatise is the search for meaningful theorems, that is, theorems which are in principle capable of empirical contradiction. The statement that a person seeks to maximize utility is (in many versions) a tautology: it is impossible to conceive of an observational phenomenon that contradicts it. The statement that demand curves have negative slopes on the other hand, is capable of empirical contradiction, and hence meaningful. There is no question but that this criterion is an excellent scientific sieve, whereby tautological and observationally undefined theorems (with which economics is well supplied) may be separated out from the theorems with content.

But it is not true that theorems about unobservable things are necessarily meaningless in a social science. On introspective grounds—and this is a perfectly valid and indeed indispensable source of information—one believes that there is such a thing as utility, and that men strive for it. One can dededuce additional conclusions (of a welfare nature) from this information, that cannot be reached by a strict positivist. Thus, we know and find it significant that, with given prices and money income, a consumer feels better off making his own choices than having someone else make them for him. The whole purpose of utility analysis in economics is to permit such welfare propositions, and such propositions are meaningful even though based upon introspection and communication.

Two different, but related, methods are employed by Samuelson to derive meaningful theorems. The first method is by use of the secondary (or sufficiency) conditions for an extremum; we may illustrate it by a very simple example. Let a competitive firm be operating at a maximum profit output where marginal cost equals price; the secondary condition is that marginal cost is increasing at this output. Let the price now rise: if the theory is correct, the output of the firm will rise or at least not fall (even if the marginal cost curve is vertical)—hence if a decline in output follows a rise in price, the theory is contradicted.

Samuelson gives a variety of illustrations of this approach (Chaps. 2–3) and then proceeds to more detailed studies of the output of the individual firm (Chap. 4) and the demand of the individual consumer (especially Chap. 5; special mention should also be made of the excellent discussion of index numbers in Chap. 6). Many of the theorems are already familiar, through the work of Slutsky, Hicks and Allen, Stackelberg, Mosak, etc., but nowhere else can one find so comprehensive a treatment.

There are only two points I would raise in connection with this first method of deriving meaningful theorems. The first is: why should not theorems on groups (industries, for example) also be derived? It is implied (p. 5) that this is impossible because an industry equilibrium does not represent a maximum position, but this is surely mistaken: the particular entrepreneurs are in the industry because it is the most profitable place to be.

My second comment is connected with the question: what if the theorem is contradicted by observation? Samuelson says it would not matter much in the case of utility theory (p. 117); I would say that it would not make the slightest difference. For there is a free variable in his system: the tastes of consumers. These tastes are not observationally (or operationally) defined, so any contradiction of a theorem derived from utility theory can always be attributed to a change of tastes, rather than to an error in the postulates or logic of the theory. The situation is less acute in the theory of the firm, for the corresponding variable (technology) is not observationally so free. But Samuelson's (and everyone's else) utility theory is still tautological as an explanation of market behavior.

Samuelson's second method of deriving meaningful theorems is to draw the consequences of dynamic stability. For example, we may assume that price will rise (fall) if the quantity demanded at the present price exceeds (is less than) the quantity supplied, and that the greater the rate of change of price, the greater the discrepancy between quantity demanded and quantity supplied (p. 263). Then if, after an initial disturbance, the price is to approach an equilibrium value, the demand curve must cut the supply curve from above. A miscellany of examples of this technique are given (Chap. 9), the most interesting being those for a simplified Keynesian system.

Thereafter Samuelson enters on a roving discussion of dynamic theory. Most discussions of economic dynamics, I feel, would better be entitled, "What I Know about Differential and Difference Equations"; with Samuelson's mathematical erudition, no doubt his version is, "What the Reader Should Know of Differential and Difference Equations." The economics, as usual, is appallingly absent. Some of the infinities of mathematical possibilities are discussed, but only in the most formal terms; there is no instance of the enlargement of our knowledge of economic processes in our society. Samuelson may reply that he is only providing tools, but who can know what tools we need unless he knows the material on which they will be used?

Although it is not an integral part of the search for meaningful theorems, welfare economics of the Hotelling variety (i.e., somewhat free of value judgments) also receives considerable attention (Chap. 8). I persist in my belief that this is one of the less fertile areas that modern economists till; it is symptomatic that we have elaborate instructions on how to form welfare judgments that do not depend on value judgments, but we have no illustration of the application of this technique to a real problem of contemporary policy. Samuelson indeed offers much support for this skeptical view, by this enumeration of the assumptions of the new welfare economics (pp. 222ff.), most of which are held to be partly invalid. But he fails to examine other facets of the problem, one of which, I think, is especially significant. When economists are writing freely on desirable policy, that is, when they are not writing on methodology, the disputes are almost always over how the economic system works, and not over the goals that should be sought. Thus, although I disagree with some of the policy recommendations in Samuelson's recent textbook, I do not differ (except in emphasis) from the goals that they are intended to subserve.

Samuelson's exposition is consistently mathematical, and presupposes a knowledge of advanced calculus, higher algebra, and differential equations (in appendices he summarizes some theorems from these branches and gives a lengthy and technical introduction to difference equations, which to my unmathematical mind seems much more difficult than say Jordan's book on finite differences). I consider this failure to provide translations for the "literary" economist a serious shortcoming of his work. He dismisses translations into words as "mental gymnastics of a peculiarly depraved type." I disagree. There is no depravity, nor is there virtue, in telling other competent economists things in a language they all can understand—there is simply responsibility to the canons of scholarship.

Sampling Inspection: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective. By the Statistical Research Group, Columbia University, Applied Mathematics Panel, Office of Scientific Research and Development. Edited by *H. A. Freeman* (Associate Professor of Statistics, Massachusetts Institute of Technology, Cambridge, Mass.); *Milton Friedman* (Associate Professor of Economics, University of Chicago, Chicago 37, Ill.); *Frederick Mosteller* (Associate Professor of Mathematical Statistics, Department of Social Relations, Harvard University, Cambridge, Mass.); and *W. Allen Wallis* (Professor of Statistics and Business Economics, University of Chicago, Chicago 37, Ill.). New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1948. Pp. xx, 395. $5.25. (London W.C.2: McGraw-Hill Publishing Co., Ltd. [Aldwych House, Aldwych]. 31*s*. 6*d*.) *Two reviews follow:*

REVIEW BY EDWIN G. OLDS

Associate Professor of Mathematics, Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania

SAMPLING inspection is the process of deciding whether to accept or reject an aggregate of material on the basis of the evidence obtained by inspecting only part of it. The evidence may be in the form of percentage of units nonconforming or may consist of measurements for each unit inspected. It is customary to denote the procedure, in the first case, as acceptance sampling by attributes and, in the second case, as acceptance sampling by variables.

Early in 1945 the Statistical Research Group, Columbia University, operating under the direction of the Applied Mathematics Panel, Office of Scientific Research and Development, was asked by the Navy to prepare a manual on sampling inspection, including tables, procedures, and principles. A manual on acceptance sampling by attributes was prepared first. The close of the war prevented completion of the work then under way on acceptance sampling by variables. The present volume is an expansion of the preliminary edition of the Navy manual.

The book has the subtitle: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective. The last four words should be emphasized since, in view of the main title, *Sampling Inspection,* the reader might otherwise expect to find some discussion of sampling inspection based on measurements. The reviewer's main criticism of this important book might be that no treatment of acceptance sampling by variables has been included. In view of the book's subtitle probably this is an unfair cavil and merely reflects the reviewer's disappointment that this companion problem still lacks the same comprehensive exposition which the present volume now makes available for acceptance sampling by attributes.

The three purposes of the book (as stated by the authors on p. 3) are: (1) "to explain the general principles of sampling inspection as a method for determining the quality and acceptability of material and as a method for controlling the quality of material"; (2) "to describe in detail a standard sampling-inspection procedure that may be readily applied to a wide variety of products"; (3) "to provide an extensive catalogue of sampling plans, usable both by those who adopt the standard procedure and by those who construct tailor-made procedures." With the restriction noted above, the book succeeds admirably in accomplishing all three purposes. It will meet the needs of three kinds of readers: those who want a ready-made sampling-inspection program, those interested in building their own programs, and students of acceptance sampling.

The book is divided into five parts. Part I consists of a chapter of introduction; Part II contains five chapters of discussion of the principles of sampling inspection for attributes; Part III, with seven chapters, gives a description of a standard sampling-inspection procedure; Part IV has four chapters on the construction of sampling tables and procedures. The foundation of the book, as noted in the preface, is the set of sampling tables presented in Part V.

As suggested in the Preface, Parts II, III, and IV have been written in such form that anyone needing a ready-made sampling inspection program can use Parts I, III, and V, omitting Parts II and IV; readers interested in building their own programs can substitute Part II for Part III; and, finally, those who want to see all of the technical details of tabular construction will find them in Part IV. Probably this plan of organization increases the scope of utilization of the book, but it requires that there be a considerable amount of repetition.

Ordinarily, the surest way to avoid pleasure from reading a novel is to read the last chapter first and certainly it is unorthodox to review a book in that order. This reviewer, however, could not resist the temptation to turn to Part V and, after much consideration, decided this would be an appropriate part to describe first.

There are four tables in Part V. Table 4, which is captioned "Sampling Plans and Operating-characteristic (OC) Curves, Classified by Sample-size Letter, and by Acceptable-quality Level (AQL) and Average Outgoing-quality Limit (AOQL)," consists of 159 pages. With a few exceptions, each page contains a single, a double, and a group sequential sampling plan together with the operating characteristic curves for each. The three plans are matched on the basis of their operating characteristic curves.

On page 288, a fairly representative page, there is a single sampling plan specifying a sample of 75 and an acceptance number of 1; a double sampling specifying a first sample of 50 and a second sample of 100, with acceptance number of 0 and rejection number of 3 for the first sample and with acceptance number of 2 and rejection number of 3 for the combined sample; a sequential plan, having seven steps of 20 units each, with decision assured at the end of the seventh sample. Printed in color, in the lower part of the page, are the operating-characteristic curves for the three plans.

The single and double plans, at least, are familiar since they appear in the Standard Sampling Inspection Tables issued by the Army Service Forces as part of ORD-M608–8. (This is not surprising since, in the words of the authors (p. vi) "the tables and procedures in this book are in large measure adapted from or based on Army Ordnance tables and procedures.") The double sampling plan, for example, is given in Table 1, Normal Lot-by-Lot Inspection, corresponding to an Acceptable Quality Level of .26 – 50% for major defects and sublots of size of 500 to 799. The approximate AOQL is recorded as 1%. In the volume under review, however, the classification is "Sample-size Letter: H; AQL class: 0.32–0.65; AOQL class: 0.90–1.5." The classification by sample-size letter, then would seem to have replaced classification on the basis of lot size. Actually the change is not so direct, or so easy to dismiss.

In order to find the appropriate sampling plan in the set of Army Ordnance Tables, the type of sampling (single or double) and the acceptable quality level must be specified; the sublot size must be known; and information regarding the level and state of control of the manufacturing process should be available. For any particular type of sampling, sublot size and acceptable quality level, three plans are available; a "normal" plan, to be used when the process average is near the acceptable quality level; a "reduced inspection" plan, when the process is in control at a level significantly better than the acceptable quality level; and a "more exacting" plan, when the average quality of inspected product is significantly worse than the acceptable quality level. The "reduced inspection" plan, when appropriate, requires about one-fifth as much sampling as the "normal" plan. The "more stringent" plan uses the same sample sizes as the "normal" plan, but lower acceptance numbers, thus causing the rejection of a greater percentage of imperfect lots. The tables contain the criteria upon which a decision between plans is made for each lot which is submitted.

The first step in abstracting the appropriate sampling plan from the set of plans provided in the book under review might be to enter Table 1 of Part V. This table is captioned "Sample-size Letter, by Inspection Level and Size of Inspection Lot." There are five "inspection levels": I, II, III, IV, V. The amount of inspection at these levels is in the ratio 1:1½:2:3:4 (approximately). A footnote states that "inspection level III will be used for most products for normal inspection," and, corresponding to this level and lot sizes 500–800 is found sample-size letter H, which appeared on page 288 in conjunction with "normal" Army Ordnance plans. Any of the other four levels, however, might be chosen as "normal," the choice depending on economic considerations.

In order to proceed further it is necessary to fix an Acceptable Quality Level. Table 2, "Summary of Sampling Plans Classified by Acceptable-quality Level and Sample Size Letter," has subtables for single-sampling, for double-sampling and for sequential-sampling. The specific plan to be adopted can be obtained here or direct reference may be made to Table 4, where three types of plans are matched as described above.

The reader who turns directly to the tables of Part V probably will be at a loss as to what inspection level to choose for "normal" inspection and, if he is familiar with other sets of sampling plans, he may wonder, perhaps, under what conditions there should be some change of plan in order to provide for "reduced" or "tightened" inspection. The term "inspection level" seems to be new. As explained on pages 45–46 and elsewhere, it is set for the product irrespective of the size of the lots into which the product is subdivided for inspection purposes. The higher the inspection level for a given lot size, the better will be the discrimination between good and bad lots. For the same acceptable quality level and lot size the plans specified for the five inspection levels, provide, in general, for a decrease in risk of accepting unwanted product with an increase in inspection level. Inspection level is not to be confused with percentage of inspection; these tables follow Army Ordnance in providing for some increase in discrimination (and, therefore, sample size) as the lot sizes increase.

Footnotes in Table 1 direct the user to drop his inspection level two steps for reduced inspection, but to retain the normal level for tightened inspection. It is unfortunate that Table 1 contains no reference to Part III, Chapter 11 (pp. 120–125) for it is here that the authors tell when and how to change inspection plans. The criteria and procedures are substantially the same as given in the Ordnance Tables.

If rejection of lots is to be interpreted as sentencing lots to 100% inspection then it is important to know the Average Outgoing-quality Limit for material inspected. Table 3, "Sample-size Letters for Average Outgoing-quality Limit (AOQL) Classes and Acceptable-quality Level (AQL) Classes," supplies the required information. It may also be used to get sampling plans when an AOQL is specified.

To indicate further the difference of these tables from the Army Ordnance Tables it seems desirable to quote the authors (p. 142).

"The major respects in which the present collection of sampling plans differs from the Army Ordnance collection are as follows: (*a*) it includes sequential-sampling plans as well as single- and double-sampling plans; (*b*) it includes plans for smaller sample sizes—for example, the single-sampling in the main Army Ordnance single-sampling table (Table C) have sample sizes varying from 75 to 1,500; in Table 2 of this book, the sample sizes vary from 5 to 1,500; (*c*) it uses a different concept of AQL; (*d*) all the plans are contained in a single table (Table 2), whereas Army Ordnance uses supplementary tables for reduced and tightened inspection and for inspection for incidental defects."

It would be very culpable for the reviewer to imply that this book contains merely a set of sampling plans together with brief instructions regarding their utilization. On the contrary, the book treats the subject of sampling inspection by attributes in great detail. The bases for the decisions necessary in installing and operating acceptance sampling are discussed from almost every angle. The methods of constructing the sampling plans are carefully illustrated. Various possible alternatives are examined and the reasons why they were discarded are given in full. The reviewer has been unable to think of any question regarding the tables which has not been answered somewhere in the book. Nor has he entertained any criticism of choice of procedures without being able to discover that the authors have anticipated the possibility of just this criticism and have prepared a strong brief in support of their procedures. This insistence on complete coverage makes the book a valuable reference, especially in view of the scarcity of general publications in the field.

The introduction to the principles of sampling inspection for attributes which is given by Part II is excellent. Using well-chosen examples the authors describe single, double and group sequential sampling; introduce the reader to operating characteristic, average outgoing quality, and average amount of inspection curves; and discuss the various factors connected with the installation and operation of sampling plans. Methods of constructing control charts for fraction defective and for defects per unit are well presented and the use of sampling plans for control sampling is discussed. Finally there is a brief, but pointed appendix on the relation of sampling inspection to design specifications.

Part III describes how the set of tables in Part V can be used as the basis of an integrated inspection program. The considerations necessary in starting a sampling procedure are outlined and then the details of the operation of a sampling procedure are set down. Step-by-step the inspector is told what to do, how to do it, and why. Here he finds rules for drawing the sample, making his decision on the lot, calculating process average, changing from normal to reduced or tightened inspection, etc. This is the part of the book which activates the tables. It is an expert job and major excerpts will, no doubt, be quoted in many an inspection manual.

The somewhat limited class of readers who are concerned with methods of constructing sampling tables will welcome Part IV with enthusiasm. Up to present time detailed knowledge of the procedures used in calculating tables has been restricted largely to those intimately connected with the actual task. Given an individual sampling plan, it is relatively simple to write the formulas and make the calculations necessary to determine how the plan can be expected to operate; given rational requirements for an individual plan, it is not too difficult to write the equations which, if they could be solved, would lead to a plan exactly meeting the specifications; but, given alternative sets of desired properties, none of which can be satisfied exactly, the problem of discovering which properties can be imposed on a set of sampling plans at a satisfactory level of uniformity and then planning the execution of the necessary computations is neither small nor easy. Therefore, anyone who has this last problem to solve has the greatest interest in the philosophy, principles and techniques basic to previous solutions. The clarity and comprehensiveness of Part IV will be appreciated.

There are a few harmless misprints in the book, and there is an occasional statement with which the reviewer could not entirely agree. On page 12, for instance, the remark is made that "the acceptance or rejection of an entire inspection lot on the basis of the quality of a sample drawn from that inspection lot provides a strong incentive for the supplier to improve the quality of his product." If, as often happens, the customer's inspection is preceded by a final inspection by the producer, then it is the improvement of this final inspection for which the customer's rejections provide a strong incentive.

In an attempt at simplification the authors have decided to redefine acceptable-quality level (AQL) as "percentage of defective items in an inspection lot such that the sampling plan will result in the acceptance of 95 per cent of submitted inspection lots containing that percentage of defective items" (p. 144). This reviewer is not alone in deploring the placing of this additional, and seemingly unnecessary, hazard in the path of the lay reader of statistics.

But, in spite of his few minor criticisms, the reviewer is convinced that this is an outstanding book. Anyone who is at all concerned with the problem of acceptance sampling by attributes needs to be thoroughly conversant with its content.

REVIEW BY PAUL S. OLMSTEAD

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

THE authors state that *Sampling Inspection* presents "in a form suitable for the peacetime needs of industry, sampling procedures for use in either acceptance inspection or process control." Only two of the seventeen chapters deal with process control. Hence, it may be stated that, basically, the book is for a consumer who is unwilling to accept any evidence concerning the quality of a lot except that given by test of a sample of that lot. The plans associated with the procedures are devised to reject approximately 5% of lots at the Acceptable Quality Level (AQL—this term is defined in a new way) and to assure delivery of product in the long run, subject to additional work by the supplier, that will meet an Average Outgoing Quality Limit (AOQL). This incidentally places a limit on the Lot Tolerance Per Cent Defective (LTPD). Plans for reduced inspection on the basis of spaced continued satisfactory quality from one supplier aim to keep rejections at AQL and acceptances at LTPD at approximately the same levels as in normal plans. Plans for tightened inspection to protect against poor quality from a supplier with a previous record showing that his quality is normally poorer than AQL keep sample sizes the same but reduce in per cent the AQL, AOQL and LTPD. This places a very severe penalty on lots at AQL in many cases, reducing the chance of acceptance from 95% to as little as 10%. No consideration is given to the possible alternative of having suppliers submit evidence of control for consecutive lots produced prior to a current lot as partial evidence for the satisfactory quality of the lot submitted. Under this scheme, test by the consumer need be only a check inspection to verify that the lot submitted conforms to what might be expected on the basis of such evidence for control.

In addition to the sampling plans (including Operating-Characteristic (OC) Curves), which constitute almost half of the book, an account is given of methods of calculation and various bases and procedures for making use of the plans. It is pointed out that these plans relate only to the "quality of a lot" expressed by "the percentage of defective items it contains," where the quality of an individual item is expressed as "defective or nondefective." Part I is a short chapter of an introductory nature. Part II defines the major concepts used in the book and discusses the principles of sampling inspection for attributes taken into account in the construction of the tables. It includes simple examples of single, double, and sequential sampling plans. Chapter 5, in particular, points out many of the difficulties associated with sampling inspection, dwelling at length on the vagaries of selecting a "random sample". The answer reached in one instance associated with spottiness is not apt to satisfy engineers who feel that one of the objects of sampling inspection is to discover evidence for just such a condition. Such evidence is often provided by keeping track of the order in which defectives are encountered. Order is not mentioned in this book. The use of sampling inspection data to obtain evidence for control is discussed briefly in Chapter 6. This chapter rightly points out that many engineers may not be satisfied with control charts based on data for which 5% of the points are expected outside control limits when the process is in fact under control. This is one consequence of the choice of definition of AQL. Incidentally, the authors use the words "take corrective action" as applying to this 5% of the points rather than the more appropriate "look for trouble." The engineer "takes corrective action" only when he finds that a criterion based on experience as associated with trouble is exceeded. In other cases, he may "look for trouble" but, ordinarily, would not find it necessary to "take corrective action." This part closes with a chapter dealing with the difference between design requirements and acceptance requirements.

Part III considers a "standard sampling inspection procedure." What is presented is a reasonably complete prospectus, with numerous examples, that may be followed as a model by those who are willing to accept the underlying principles mentioned above. Most engineers will find much of the material useful as a general guide to a larger number of things that may cause trouble in setting up inspection procedures. Part IV discusses the coverage of the tables and the methods used in computation, dealing particularly with shortcuts, limits of error, and approximations. Some mathematical statisticians may object to the use of the Poisson Exponential as an approximation to the Hypergeometric Series. Considerable space is devoted to showing how to compute the average amount of inspection (ASN—Average Sample Number) required to reach a decision to accept or reject for each of the three types of sampling. Readers should be careful not to confuse this average amount of inspection with the average amount of inspection required to give assurance that the quality of product meets the AOQL. The latter takes into account the detailing of rejected lots and therefore may differ but little for the three types of sampling. Comparative values of ASN are given for only a few of the sampling plans.

Part V is devoted to presentation of the tables. These, in most cases, show single, double and limited step curtailed sequential sampling plans for ranges in lot size (cross referenced to a sample size letter) and ranges in AQL and AOQL. On each page along with a table showing three sampling plans is a chart with the OC curves for these plans. These curves, being in color, are subject to printing errors as explained on page x. The book also contains a consolidated list of references, a glossary, and an index. The relatively few errors, mainly associated with printing, should not cause the reader difficulty.

It is quite likely that many engineers will find the sampling tables useful in their work. The underlying philosophy associated with their use, however, may lead to a somewhat different "standard sampling inspection procedure" —one that would place more emphasis on process control and less on acceptance inspection than the one outlined in Part III of the book. Consequently, this reviewer would classify *Sampling Inspection* as a good reference book but not a textbook.

How to Chart: Facts From Figures With Graphs. *Walter E. Weld* (Investment Adviser, 227 Rawson Road, Brookline, Mass.). Preface by *A. C. Haskell* (Codex Book Co., Inc., Norwood, Mass.). Norwood, Mass.: Codex Book Co., Inc. (74 Broadway), 1947. Pp. xiv, 218. $3.00.

REVIEW BY KENNETH W. HAEMER
Statistician, American Telephone and Telegraph Company
195 Broadway, New York 7, New York

THIS is a book for the businessman who knows only a little about charts or about how to make them. On the other hand, several of the seventeen chapters discuss chart uses and methods that even the experienced chartist—businessman or not—might employ with profit.

Item number one on the credit side is the author's style of writing. It is easy to read, free of unnecessary technical terms, and effective in conveying the author's ideas. It can be summarized as conversational rather than academic.

The illustrations are excellent, plentiful, and well drawn; they accurately illustrate the point under discussion, a feature more often absent than present in books about charts. It is a sad paradox that so many books on graphics are so poorly illustrated: Mr. Weld is to be congratulated on refusing to follow this "do as I say—don't do as I do" line of least resistance.

Throughout the book, the charts are explained in terms of specific business applications. Although this approach has the mild disadvantage of seeming to restrict the chart to only one type of use or application, it does succeed in bringing the charts to life and in pointing up exactly how each fits a particular type of business activity. It should be an effective means of introducing charts to any businessman who is just beginning to discover the extent of their usefulness.

The principles of chart presentation are skillfully explained. It is perhaps regrettable that Mr. Weld does not accept the principle that simple bar comparisons (1-scale measurement) are generally best shown as horizontal bars, and that the vertical arrangement is best reserved for 2-scale bar presentation. Also, in his discussion of multiple-vertical-scale presentation he has mentioned only the advantages and has neglected to point out the considerable disadvantages. In almost all other respects, however, the graphic principles set forth are beyond reproach.

Considerable space is given to instruction on how to prepare charts. Although these instructions naturally reflect the author's own experience, and occasionally ignore the fact that there are alternatives, they should be extremely helpful to the beginning chart maker. Some of these practical hints are especially valuable because they so seldom find their way into print. For example, Mr. Weld points out that since the color light blue does not photograph, lines and other markings that are *not* to appear on the reproduced copies should be made on the original with a light blue pencil. This may seem to be a small point, but it is one that saves troubles and dollars, and one that incidentally demonstrates the author's practical point of view.

This book focuses attention on three types of charts that in many other books are discussed only briefly or are not mentioned at all. Each of them is a type developed from a special kind of scale: of the three, the semi-logarithmic chart properly receives the most emphasis. However, the other two, probability scale and reciprocal scale charts, are discussed at sufficient length to explain how and when they can be used to advantage.

The first of the two chapters on the semi-logarithmic chart provides an excellent introduction to the mechanics of this form of presentation. Of course, it is possible to quibble with some of the things the author has to say about the semi-logarithmic chart: it is a subject that can be quibbled about with very little effort. For example, he says " . . . it is often desirable and extremely helpful to know the *rate* of change, or percentage of gain or loss. The kind of chart which gives this information is called a ratio (or semi-logarithmic) chart." This reviewer holds to the belief that a semi-log chart shows only *relative* rate of change. But the preciseness of the term is indeed a minor matter and does not detract from the general soundness of the exposition. In Chapter 7, which discusses the uses of semi-log charts, Mr. Weld not only explains when the type of presentation is appropriate, but repeatedly points out that it should not be used with the mistaken idea that it can simultaneously serve the purpose of an arithmetic chart. This one point alone, if thoroughly absorbed by the reader, would be worth the price of the entire book.

The author's discussion of the probability scale and its uses in business charting (Chap. 8) is easy to read and is informative. Here, as elsewhere throughout the book, the statistical propriety of the applications cited, or the conclusions concerning these uses, are not considered to be within the scope of this review: Judged solely as an explanation of its function as a tool of graphic analysis, this chapter is a good introduction to this type of chart. Although the reciprocal chart is primarily a means of computation rather than of presentation, Mr. Weld makes a good case for devoting a chapter to it (Chap. 9). His explanation of its wide usefulness should encourage a more extensive use of this little-known chart.

In this reviewer's opinion, two other chapters deserve special mention. Chapter 15, "Charts of Various Kinds," is a quick review of assorted chart varieties—from progress charts to trilinear charts—which the author feels warrant inclusion. This chapter provides a useful, if limited, supplementary "idea" file. Chapter 16, "Exhibits Good and Bad," is a chart clinic in which seventeen selected examples are subjected to critical examination. These should be two of the most valuable chapters in the book for the novice in chart presentation.

Since it is inevitable that author and reviewer seldom agree on all things, some chips must fall on the unfavorable side.

To this reviewer, the emphasis in the book seems faulty. It appears to be designed to focus attention on the major interests of the author rather than on the probable interests of the reader. Because of the author's very special field of interest, these two are unlikely to be the same.

For example, stock market analysis charting and business forecasting fill three chapters and parts of others. This is two chapters more than is allowed to any other subject except semi-logarithmic charts, and seems out of proportion in an introductory discussion to business charting. Indeed, this ardent exposition seems appropriate for only one group of businessmen—investment analysts—among which number the author presumably is one. Although other businessmen readers will no doubt be intrigued to learn that there are stock market patterns known as "The Head & Shoulders," and "The Rounding Bottom," to mention two of many, few are likely to feel an overwhelming desire to put this special knowledge into use.

On the other hand, the subject of map presentation is covered with one illustration and a few sentences. So fragmentary a treatment of so important a subject seems scarcely worth including.

In several spots, the sequence seems unnatural. Stock market charting (Chaps. 10–11) and Business Forecasting (Chap. 13), are separated by a chapter on "What Should Be Charted to Help the Business Manager." Chapter 14, "Making and Reproducing Graphs," is sandwiched between "Business Forecasting" and "Charts of Various Kinds."

There is a considerable amount of overlapping. This need not be a fault if it is carefully planned with a legitimate purpose, but in this case much of it has the appearance of being partly incidental and partly accidental. For example, one of the subheads under Chapter 17 is "Procuring Suitable Chart Sheets." Under Chapter 14 there is a subhead "Selection of Suitable Chart Paper." In fact, the Codex Company or its products are mentioned one or more times on 18 pages throughout the book. It is only fair to point out that several of these mentions include invitations to call on the Codex Company for further advice or guidance. Nevertheless, this reader found it disturbing to have the publisher's name thrust forward with such artless frequency. Another example of overlapping: One of the items covered in Chapter 4 is the matter of choosing lines or spaces to represent units of time. Although no mention is made in the chapter heading, Chapter 5 also discusses this problem at some length. There is, of course, some advantage in repetition at strategic points, but this reader for one prefers to find most of the information on a given subject set down on adjacent pages.

Although it makes for ease of reading and a fast moving exposition, the author's penchant for singling out one device or one method and neglecting to mention the alternatives, does a disservice to the reader who wants to know the best way rather than just a good way. For example, when discussing lettering guides, Mr. Weld mentions only Wrico lettering guides. In spite of the fact that this reviewer leans toward Wrico guides, it would seem desirable to tell the reader at least what others exist, and in order to be of maximum help to explain the advantages and disadvantages of each.

This reviewer is of the opinion that no one book about charts can hope to answer all chart questions. The design and use of charts involves so many decisions based on judgment and on knowledge outside the chart itself, that any book, however extensive, must leave many questions unanswered. The author of a book on charts is confronted with three major choices: a thorough job on principles, applying to all fields of charting; a thorough job on the detailed procedures for a limited field of charting; or a compromise combining some of each. In his present book, Mr. Weld has chosen to follow the latter course. He has succeeded very well in presenting the principles of simple charting, spiced with enough specific applications to make them appetizing and easy to digest. Although it is directed primarily to businessmen, this book has something to say also to nonbusiness chart makers.

In addition to providing a helpful introduction to the beginner, this book fits neatly onto the bookshelf, next to more voluminous works, where it should serve as a compact reference to the author's special field of interest.

NOTE: As a result of the reduction in the number of pages in this issue of the *Journal*, the Statistical Methodology Index is being omitted. Readers are invited to write to the Secretary's office concerning the desirability of continuing this series of bibliographies of the literature on statistical methodology.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

CONTENTS OF VOLUME 43

The 107th Annual Meeting

Program. 299

Minutes of the Annual Business Meeting 308

Articles 1, 211, 357, 497

Book Reviews 165, 312, 461, 635

Letters About Books 172

Statistical Methodology Index 174, 356, 500, 668

Index to Volume 43, 1948

Articles, by Author 620

Book Reviews, by Author 621

Reports and Official Notices

Report of the Board of Directors 311

Report of the Secretary 314

Report of the Treasurer 314

Report of the Auditors 315

Committees and Representatives for 1947 307

INDEX TO VOLUME 43: 1948

ARTICLES

Arnold, David R., and Spurr, William A., *A Short-Cut Method of Fitting a Logistic Curve* 127

Bowker, Albert H., *A Test for Symmetry in Contingency Tables* . . . 572

*Brown, Weir M., *Measuring Physical Inventories* 377

*Campbell, George C., *Problems with Sampling Procedures for Reserve Valuations* 413

*Cody, Donald D., *Sampling Errors in Mortality and Other Statistics in Life Insurance* 442

Deemer, Walter L., Jr., *The Use of Mark Sensing in a Large Scale Testing Program* 40

Dixon, W. J., and Mood, A. M., *A Method for Obtaining and Analyzing Sensitivity Data* 109

*Epstein, Benjamin, *Application of the Theory of Extreme Values in Fracture Problems* 403

Finney, D. J., *Main Effects and Interactions* 566

*Greville, Thomas N. E., *Recent Developments in Graduation and Interpolation* 428

Grubbs, Frank E., *On Estimating Precision of Measuring Instruments and Product Variability* 243

Correction 564

Harris, Marilyn; Horvitz, D. G., and Mood, A. M., *On the Determination of Sample Sizes in Designing Experiments* 391

Hochstim, Joseph R., and Stock, J. Stevens, *Commercial Uses of Sampling* 509

Horvitz, D. G.; Harris, Marilyn, and Mood, A. M., *On the Determination of Sample Sizes in Designing Experiments* 391

Kirchen, C. J., *Levels of Significance for Variance Ratio of Two Samples of Equal Size* 564

Lasken, Herman, *Direct Determination of Compass Settings for Proportionate Area Pie-Charts* . . . . John V. Spielmans 94

Lebergott, Stanley, *Earnings of Nonfarm Employees in the U.S., 1890–1946* 74

*Lew, Edward A., *Mortality Statistics for Life Insurance Underwriting* . . 274

Mood, A. M., and Dixon, W. J., *A Method for Obtaining and Analyzing Sensitivity Data* 109

*McDiarmid, Fergus J., *Life Company Investments and the Capital Markets* 265

*Miller, Morton D., *Statistical Problems in Developing Voluntary Medical Expense Insurance Plans* 290

Mood, A. M.; Harris, Marilyn, and Horvitz, D. G., *On the Determination of Sample Sizes in Designing Experiments* 391

Mosteller, Frederick, *On Pooling Data* 231

Niessen, Abraham M., *Actuarial Estimates for Public Sickness Insurance Plans* 61

Novick, David, and Steiner, George A., *The War Production Board's Statistical Reporting Experience* 201, 463, 575

*Paradiso, Louis J., *Significance of Inventories in the Current Economic Situation* . . . 361

Richardson, Lewis F., *Variation of the Frequency of Fatal Quarrels with Magnitude* . . . 523

Schuyler, Garret L. *The Ordering of* n *Items Assigned to* k *Rank Categories by Votes of* m *Individuals* . . . 559

*Snedecor, George W., *A Proposed Basic Course in Statistics* . . . 53

Spielmans, John V., *Profile Graphs* . . . 96

Spurr, William A., and Arnold, David R., *A Short-Cut Method of Fitting a Logistic Curve* . . . 127

Stephan, Frederick F., *History of the Uses of Modern Sampling Procedures* 12

Stock, J. Stevens, and Hochstim, Joseph R., *Commercial Uses of Sampling* . . . 509

*Thorp, Willard L., *Statistics and Foreign Policy* . . . 1

Thurstone, L. L., *The Edge-Marking Method of Analyzing Data* . . . 451

Warburton, Clark, *Bank Reserves and Business Fluctuations* . . . 547

BOOK REVIEWS

Abay, Gyula, *Bevezetés a Statisztika Tudományába: Rész I* . Eugene Lukacs 135

Albert, Adrian A., *College Algebra* . . . Paul S. Dwyer 136

American Society for Quality Control, First Annual Convention and Second Midwest Quality Control Conference, June 5–6, 1947, Conference Papers . . . H. A. Freeman 353
. . . Eugene L. Grant 354

Burton, Clement, *A Concise Manual of Statistics: With Special Reference to the Requirements of Students for Municipal Examination, Second Edition* . . . Harry Pelle Hartkemeier 137

Chapin, F. Stuart, *Experimental Designs in Sociological Research* . . . O. Kempthorne 489

Columbia University Statistical Research Group, Applied Mathematics Panel, Office of Scientific Research and Development, Edited by H. A. Freeman, *Sampling Inspection: Principles, Procedures and Tables for Single, Double and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective* . . . Edwin G. Olds 605
. . . Paul S. Olmstead 611

Dahlberg, Gunnar, *Mathematische Erblichkeitsanalyse von Populationen* . . . Oystein Ore 139

Davies, Owen L., editor. *Statistical Methods in Research and Production: With Special Reference to the Chemical Industry* . . . D. B. DeLury and A. E. R. Westman 318
. . . E. L. Lehmann 320

Dewey, Edward R., and Dakin, Edwin F., *Cycles: The Science of Prediction* . . . Milton Friedman 139
. . . Max Sasuly 141

Eisenhart, Churchill; Hastay, Millard W., and Wallis, W. Allen, editors. Statistical Research Group, Columbia University, *Selected Techniques of Statistical Analysis: For Scientific and Industrial Research* *and Production and Management Engineering* . . . Henry Scheffé 321

Finney, D. J., *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve* . . . . Margaret Merrell 145

. . . . Joseph Berkson 148

Fisher, Ronald A., *The Design of Experiments, Fourth Edition* . . .

. . . . A. E. Brandt 323

Florence, P. Sargant, *Investment, Location, and Size of Plant: A Realistic Inquiry into the Structure of British and American Industries* . . .

. . . . Edgar M. Hoover 601

Furlan, L. V., *Das Harmoniegesetz der Statistik: Eine Untersuchung über die metrische Interdependenz der sozialen Erscheinungen* . . . . . .

. . . . Hilda Geiringer 325

Guerreiro, Amaro D., *Manuel de Estatistica* . . . Hugo Muench 153

Hastay, Millard W.; Eisenhart, Churchill, and Wallis, W. Allen, editors. Statistical Research Group, Columbia University, *Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering* . . . Henry Scheffé 321

Hays, Samuel, *An Outline of Statistics, Third Edition* Harold Nisselson 155

. . . . Edgar Z. Palmer 156

Heidingsfield, Myron S., and Blankenship, Albert B., *Market and Marketing Analysis* . . . . Lester R. Frankel 157

. . . . Philip J. McCarthy 160

Hill, A. Bradford, *Principles of Medical Statistics. Fourth Edition* . .

. . . . J. Yerushlmy 492

Institute of Mathematical Statistics Committee on the Teaching of Statistics, Report of the, *The Teaching of Statistics* . . . .

. . . . .Truman L. Kelley 493

. . . . Alan E. Treloar 597

Kelley, Truman Lee, *Fundamentals of Statistics* . . . Cyril Burt 328

. . . . Horace W. Norton 330

Kenney, John F., *Mathematics of Statistics*, Part One, Second Edition . .

. . . . Howard Levene 332

. . . . Ronald W. Shephard 334

Levens, A. S., *Nomography* . . . . M. Kac 496

Mortara, Giorgio, *O Custo de Produção do Homem Adulto e Sua Variação em Relação à Mortalidade* . . . . T. N. E. Greville 161

Peatman, John Gray, *Descriptive and Sampling Statistics* . . . . .

. . . . Leon Festinger 335

. . . . Frederick F. Stephan 337

Royal Statistical Society's Committee on the Teaching of Statistics, *The Teaching of Statistics in Universities and Colleges* Truman L. Kelley 493

. . . . Alan E. Treloar 597

Samuelson, Paul Anthony, *Foundations of Economics Analysis* . . .

. . . . George J. Stigler 603

. . . . Gerhard Tintner 497

Scarborough, J. B., and Wagner, R. W. *Fundamentals of Statistics* . .

. . . . Leo Aroian 339

Schweng, Lorand, *Statisztika Módszertani Alapvetés* . Eugene Lukacs 135

Smart, L. Edwin, and Arnold, Sam, *Practical Rules for Graphic Presentation of Business Statistics* . . . . Rudolf Modley 341

Smith, Edward S., *Control Charts: An Introduction to Statistical Quality Control* . . . . Joseph M. Juran 343
. . . . Mason E. Wescott 343
Thompson G. Clark, *Forescasting Sales* . . . . Charles F. Roos 163
. . . . H. D. Wolfe 165
Thionet, Pierre, *Méthodes Statistiques Modernes des Administrations Fédérales aux États-Unis* . . . . . . W. Edwards Deming 346
Weld, Walter E., *How to Chart: Facts From Figures With Graphs* . . . . Kenneth W. Haemer 613
Westman, A. E. R., *An Extension Course in Statistical Quality Control* . . . . A. C. Cohen, Jr. 350
. . . . Paul Peach 351
Wright, Wilson, *Forecasting for Profit: A Technique for Business Management* . . . . William A. Spurr 167
. . . . Victor von Szeliski 168
Zetterlund, Stig, *Toleranser och Passninger*. . . . Cecil C. Craig 352